Using population genetics to assess the dispersal patterns of the New Zealand mayfly *Coloburiscus humeralis* in a landscape context.

A thesis

submitted in partial fulfilment

of the requirements for the degree

of

Master of Science in Zoology

at the

University of Canterbury,

New Zealand

By Matthew F. Wallace

University of Canterbury

2013

Abstract

Mayfly species, such as *Coloburiscus humeralis*, are an important organism in freshwater ecosystems. As well as often being crucial intermediary links in food webs, mayfly nymphs provide a number of vital in stream functions that contribute to ecosystem health. The threat to stream invertebrate populations from habitat destruction as well as chemical and nutrient additions from agricultural land use has had considerable attention in New Zealand. However, less has been done on the potential barriers to dispersal that may negatively affect the health of stream insect populations. The goal of this study was to use AFLP techniques to measure the genetic structure of populations between streams and across catchments, and thereby to better understand the dispersal of this species across the landscape. I sampled populations in 10 forested streams, ranging from populations separated by <2km to 780 km. I found that in the Arthur's Pass study area, populations of *C. humeralis* had high levels of genetic divergence even between the closest streams in the study (Φ_{PT} = 0.065, p = 0.011). In order to identify the landscape features that either constrain or assist mayfly dispersal, I employed a resistance modelling technique based on electrical circuit theory to simulate how particular landscape factors contribute to the observed genetic structure. My resistance modelling results suggest that open areas provide barriers to dispersal while gene-flow was more likely to occur when streams are separated by connected native forest.

Acknowledgements

Firstly I would like to thank my supervisors Jon Harding and Marie Hale for the amazing support and advice throughout. Secondly I would like to thank my parents Brian and Bridget Wallace for their constant support and encouragement. A very special thank you to Maggie Tisch for the laboratory assistance, as well as all the encouragement and advice during the period when I was unable to produce genotype worthy of mention. A huge thank you to the Molecular Ecology Research Group, especially Michael Bartlett, Rachel Van Heugton, Ilina Cubrinovska, and Alex Rodriguez for all advice, great conversation and generally making the lab a pleasant place to be. Thank you to everyone in F.E.R.G. for all their support and advice on all things freshwater. Much love to my little girl Georgia, who kept me sane and happy during my most stressed times in this process. Finally, this research would not have been possible without the support of the Brian Mason Trust.

Table of Contents

Chapter 1

Introduction

Rationale

The ability of organisms to disperse is an essential attribute driving the structure and composition of biological communities. Ronce (2007) defines dispersal as any movement by individuals or propagules that leads to gene flow across space. This emphasis on gene flow is important because the effects of dispersal are far ranging, not only is it fundamental to the fitness of individuals but to the genetics and dynamics of populations as well as the distribution of species (Bowler and Benton, 2004). Therefore understanding why organisms disperse and the costs that dispersal entails is essential to the study of populations.

The high cost of dispersal

The importance of dispersal to the success of a species over evolutionary timescales can be demonstrated by the diverse array of strategies and morphologies that have evolved to facilitate this process. Even within the same population dispersal strategies can vary (Dytham, 2009). For example, in the western bluebird, it is the more aggressive individuals that are more likely to occupy the edges of a species range as they are more likely to succeed in a new territory, while less aggressive birds are more often found in the centre of the species range where social tolerance

permits higher population density (Duckworth, 2008). The development of dispersal adaptations can often require substantial investment of resources by the individual and tradeoffs with other life history traits (Zera and Denno, 1997).

The evolution of powered flight in insects is recognised as an important factor in their success at colonising much of the planet and their high diversity (Dudley, 2002). However, some insect species have both a winged and an apterous form and this flight polymorphism is relatively common, which may indicate that there are complex tradeoffs between dispersal ability and other life history traits (Roff, 1986). In particular, various studies have shown that the apterous forms often have significantly higher fecundity than their winged conspecifics and that the reproduction of winged individuals is often delayed substantially compared to wingless morphs (Zerra and Denno, 1997; Roff, 1986). Therefore, despite the high cost of wings and their musculature, the evolutionary importance of this dispersal mechanism means that it's continued even at the cost of high fecundity.

There is little debate in the literature that dispersal is a costly exercise (Rousset et al, 2009). During a dispersal event the individual may have to cross an area that it is not adapted to (e.g., a mountain or human altered landscapes), and in many cases overcome increased risk of mortality in order to find another suitable habitat (Ronce, 2007). In a study on the root vole (*Microtus oeconomus),* it was shown the vole has risky dispersal behaviours in the face of rising population density that led predatory bird species to have a strong influence on vole population dynamics (Ims and Andreassen, 2000). Additionally, the possibility of not finding a suitable habitat or mate makes the dispersal stage particularly dangerous for the individual (Gu et al., 2006).

When an individual does find a suitable habitat to colonise, it may often have to overcome other potential challenges to its ability to spread its genes in the new habitat. For example, the individual may be at a distinct fitness disadvantage compared to those already occupying the habitat, because resident species may have evolved adaptation to conditions in that habitat across multiple generations (Bowler and Benton, 2004). Dispersal may be indirectly selected against because immigrants are less likely to hold locally advantageous alleles (Ronce, 2007). Despite their high cost, dispersal adaptations and strategies are often remarkably stable over evolutionary timescales. For example, the relatively simple but extremely effective direct flight mechanism of modern dragonflies and damselflies (Insecta: Odonata) is remarkably similar to the body plan of Protodonata found in 300 million year old Carboniferous deposits (Carpenter and Burnham, 1985). Therefore, if the ability to disperse effectively is an evolutionarily stable trait across such a range of species, the evolutionary benefits must be great.

Leaving home: Why organisms disperse

When examining why organisms disperse, it is important to consider both the mechanisms which induce individuals to engage in such a risky behaviour but also how dispersal related alleles are maintained at the population level. While the necessity to disperse can lead to a high mortality rate, the benefits to the wider population mean that dispersal adaptations are retained over long timescales. In other words we need to consider both the ultimate and proximate causes of dispersal.

Variation in an individual's fitness in multiple habitats is a key driver in the evolution of dispersal. However, this can be unpacked further to be described as kin interactions, the avoidance of inbreeding and habitat heterogeneity in general (Bowler and Benton, 2004). Early modelling work by Hamilton (1964) suggests that kin selection will favour dispersal if it acts to reduce sibling competition at the site of their birth. While separating kin competition and inbreeding avoidance is difficult empirically, Moore et al. (2006) used the male fig wasp (*Platyscapa awekei)* to demonstrate its importance to dispersal in real populations. Since these males are extremely inbred haplo-diploids, they are unlikely to suffer any appreciable inbreeding costs. While males exposed to low levels of kin competition were unlikely to leave the natal fig, those males found in figs with high densities of related individuals were significantly more likely to disperse. In their modelling paper, Poethke et al. (2010) showed that kin selection could be more important to dispersal rates than individual selection. In particular, they found that models that incorporated kin interactions could account for a ten times higher rate of optimal dispersal compared to models that looked at individual selection alone (Poethke et al., 2010).

The avoidance of inbreeding depression has long been posited as a central force driving individuals to disperse (Charlesworth and Charlesworth, 1987). To avoid the serious threat to fitness that increases in the expression of recessive deleterious alleles caused by incestuous mating (often as the result of small population sizes), it would appear that dispersing away from related individuals would be an evolutionarily beneficial practice. Using a game theory approach to modelling this relationship, Perrin and Mazalov (1999) found that when assuming that inbreeding avoidance is the only reason for dispersal, the outcome is always complete

philopatry for one sex and a dispersal rate dependent on patch size and mating strategy for the other. While this model does not take into account the complexity of natural populations, it is interesting to note that many species, especially of mammals and birds, have a highly philopatric sex with the other more likely to disperse (Pusey et al, 1987). This is not always the case however, and in a study of the co-operative breeding pied babbler (*Turdoides bicolor),* it is not sex bias but a more general dispersal distance mechanism decreasing kinship mating (Nelson-Flower et al., 2011). This is most likely due to the shared rearing of chicks in this species, and thus coincides with predictions of earlier models discussed earlier (Hamilton, 1964; Poethke et al, 2010; Perrin and Masalov, 1999).

Natural populations reside in environments that display varying degrees of spatial and temporal heterogeneity and because of this, individuals may face a variety of reasons both to disperse and not to disperse, thus predicting evolutionary dynamics has been elusive (North et al., 2011). In heterogeneous environments, good habitats can occur, but organisms may also migrate away from these optimal patches. Theoretical studies suggest that in a heterogeneous environment, the costs of dispersing may outweigh the benefits (Hastings, 1983; Holt, 1985). More recent studies that take into account the properties of unstable populations, like demographic stochasticity, have shown it to be possible that dispersal may increase with spatial heterogeneity (Cadet et al. 2003; Parvinen et al. 2003).

Variation in availability of resources over time has long been proposed as a strong driver of the evolution of dispersal. Gadgil et al., (1971) showed that in a temporally variable environment dispersal may evolve over many generations as a bet hedging strategy to enable a population to take advantage of the full potential resources of the

area. The variation in persistence of ephemeral habitat patches over time has been shown to drive the ratio of dispersal strategies in a number of insect species. For example, Denno et al., (1996), demonstrated that in the polymorphic planthopper *Prokelisia marginata*, there was a highly significant negative relationship between the percentage of winged individuals and the persistence of habitat, when there was a decrease in the quality of the plant the population was living on the population would respond by producing more winged individuals.

As well as these ultimate causes for dispersal, there often exist complex tradeoffs between dispersing and staying close to the location of one's birth. For many species dispersal is a plastic trait, with individuals responding to complex environmental cues when deciding whether to disperse (Clobert et al., 2009). Theoretical models have suggested that the most evolutionarily stable dispersal strategy in plants is a plastic one, because as the population ages the most effective seed dispersal strategy is also likely to change (Ronce et al., 2005). In an experimental study on the plant *Crepis sancta,* it was found that with increasing stress the proportion of seeds with enhanced dispersal capabilities was significantly increased (Imbert and Ronce, 2001).

Dispersal is often a response to increased density (Bowler and Benton, 2004). Individuals that find themselves in an increasingly crowded population are likely to suffer a decrease in fitness and therefore it will be in their direct interests to disperse further afield. In the lizard, *Lacerta vivipara*, the propensity to disperse is governed by the social tolerance of individuals within the population (Cote and Clobert, 2007). While some animals were highly tolerant of being in close proximity of conspecifics, the individuals most likely to disperse were those found to be socially less tolerant (Cote and Clobert, 2007). The maintenance of this behavioural polymorphism over

time allows the population to both fully exploit the patch that they currently inhabit but also means that competition is continuously reduced and overtime the population has an outlet to fill empty patches in the surroundings.

There are several other factors that act as proximate causes for dispersal. It is not surprising that empirical studies have shown that in an area that has experienced a reduction in the food resource, the dispersal rate will often rise (Schneider, Dover & Fry, 2003). Increases in the local abundance of predators and parasites will also give individuals within the population impetus to find less dangerous areas to make a living. Experimental work with the pea aphid (*Acyrthosiphon pisum*) showed that the population responded to the presence of a ladybug predator by increasing the number of winged individuals in the population (Sloggett & Weisser, 2002). The nature of various mating strategies means that many individuals have very limited opportunities to mate, moving away from their natal habitat is often their only opportunity to reproduce. In fact, this has been proposed as one of the leading mechanisms for how dispersal rates can remain high even from good habitats (Leturque and Rousset, 2003). While each one of these factors is likely to influence dispersal rates, it is important to realise that in natural populations these forces will likely interact in dynamic ways in response to fluctuations in a range of ecosystem and environmental factors. Therefore, modelling and empirical studies that wish to develop the most reliable and predictive analyses of dispersal patterns, should look to accommodate as many of these factors as possible.

Anthropogenic influences on dispersal

Human mediated changes, such as deforestation and habitat destruction have meant that populations are increasingly separated and fragmented and a there is a general widespread decline in the availability of suitable habitats (Tillman et al., 2001). In a study of European grasslands, this decrease in habitat availability has been connected to both an initial loss in biodiversity and also a delayed loss in species, referred to as an extinction debt (Krauss, et al., 2010).

At a larger scale, the human mediated movement of species around the globe has resulted in the emergence of a few 'winners' at the expense of a much larger number of 'losers'. For example, the introduction of invasive species has led to a homogenisation of ecosystems previously rich in species (Clavel et al., 2010). In islands like New Zealand the impact can be particularly extreme, with many locally adapted species not having evolved phenotypes to successfully co-exist with the invader (Sax and Brown, 2000). The introduction of large predatory salmonids into New Zealand rivers and streams has resulted in increased pressure on native galaxiids and benthic invertebrates, such as mayflies (Townsend, 2003). By reducing the number of habitats available for endemic species, the introduction of exotic invaders can lead to the fragmentation of previously continuous populations, meaning that dispersers will have to travel further to find habitat in which their offspring will have a reasonable chance at survival.

All of these forces can interact in non-linear ways which can a result in a lowering of the resilience of species in an ecosystem. This can potentially culminate in sudden and unpredictable catastrophic collapse; with all of its associated ecological and economic repercussions (Scheffer et al, 2001). For example, gradual changes in the

agricultural watershed of Lake Apopka (Florida, USA) were attributable to eutrophication, but a 1947 hurricane wiped out aquatic plants which led to a more permanent collapse in water quality (Scheffer et al, 2001). Furthermore, many ecosystems can be regarded as a network of locally interacting patches that rely on emigration to remain viable (Hanski, 1999). Freshwater ecosystems are very good candidates for consideration as meta-communities as they are comprised of discrete patches embedded within a landscape comprised of unsuitable habitat (Bohonok et al., 2003). If the existence of a habitat patch relies upon a constant stream of migrants, then the loss of intervening patches due to one or more of the above anthropogenic pressures may result in dispersal barriers. The underlying complexity of the effects of fragmentation can mean that determining the species likely to be affected can be difficult (Ewers and Didham, 2006). For example, in fragmented landscapes, being a rare but effective disperser can trump being a competitively dominant but poor disperser. In tropical beetle assemblages common species were more likely to go locally extinct in small habitat fragments than their rarer counterparts (Didham et al, 1998). Furthermore, increasing isolation can lead to a number of genetic and demographic pressures that can hasten the local extinction of the species (Bohonok et al., 2003).

What are the costs and benefits of dispersal for aquatic insects?

While freshwater insects face many similar drivers to disperse as terrestrial taxa, those that live in streams, have additional challenges. Organisms that live in moving water have to deal with flow, streams act as conveyor belts, moving energy and material (e.g. gravels, sediments, detritus and prey) continuously downstream.

Gibbons et al., (2010) showed that during the time an aquatic juvenile is present in the stream a large proportion of the substrate may be moved considerable distances. Also, the frequency of floods is a cause of downstream movement, at least for large-scale 'catastrophic dispersal events' (Bond & Downes, 2003). For example, trichopteran larvae have been moved 670m downstream after a significant flood event (Neves, 1979). The distance moved downstream due to drift can be extremely variable. McLay (1970) found that following disturbance of the substrate mean distance varied from 0.5m to 19.3m. However, one study suggests that downstream dispersal may not be as extensive as previously believed (Lancaster et al, 2011). They found that in the mayfly *Baetis rhodani,* most neonate and mid-sized nymphs stayed close to the natal riffle, while the largest instars did not disperse nearly as far as commonly assumed. Therefore population models that assume that density variations along the stream channel are simply a snapshot in time and that mixing and high levels of downstream dispersal are the dominant forces in streams (Downes & Reich 2008) may not be as widely applicable as previously thought. However, New Zealand streams are characterised by high levels of flow variability, and therefore downstream dispersal may play a greater role in the population dynamics of benthic invertebrates than the more stable British streams of the Lancaster et al. (2011) study.

For stream invertebrates, the cost of aerial dispersal is probably high. For the insect orders that are found predominately in moving waters (Ephemeroptera or mayflies, Plecoptera or stoneflies, and Trichoptera or caddisflies), the majority of their life-cycle is spent as a larva or nymph on the benthos of the stream, while in many cases some adult stages (e.g. mayflies) do not feed and are very short lived (Hinton, 1948). Therefore the adult stage is almost exclusively required for mating and

dispersal, and this may give us insight into the evolutionary importance of dispersal to aquatic insects. Furthermore, freshwater environments can be ephemeral, as streams may periodically dry due to climatic conditions and water abstraction (Bilton et al. 2001), and it may be that without a constant supply of individuals from adjacent streams that local extirpations may be frequent. Changes predicted in the climate for the 21st century (Fung et al, 2013), imply that conditions instreams and rivers will also become more variable, increasing the importance of dispersal for the long-term health of aquatic species.

By their nature, streams represent spatially discrete ecosystems. In order to move to an adjacent watershed, dispersing insects may have to negotiate terrain that is both dangerous and complex, increasing the risk of mortality. Briers et al., (2002) found that 90% of adult stoneflies were captured within 11 metres of the stream edge (which may indicate a low requirement for lateral dispersal for stoneflies). However, because many of these studies rely on limited trapping techniques to capture adult dispersers, it is unlikely that rare but potentially important long distance lateral dispersers will be adequately sampled (Petersen et al., 2004). It would therefore appear that, as well as the more general costs and benefits mentioned previously, this and the regulating influence of flow on stream ecosystems are important factors in determining the dispersal propensity in stream invertebrates.

The consequences of dispersal limitation

In order to manage species and ecosystems, conservation managers need reliable data to inform them of the most pressing threats. By viewing a species as populations embedded within a heterogeneous and spatially diverse landscape, many of the challenges to the long-term success of a species can be put into focus. The consequences of the breakdown in connectivity between habitats for a species can put the population at increased risk of losing their resilience to a multitude of other stressors (Scheffer et al, 2001).

With increasing levels of habitat fragmentation it is also likely that the risks of an individual dispersing across fragmented landscape will increase. The ability to disperse between suitable patches will be limited by the distance between these patches. Thomson and Townsend (2006) found that for freshwater invertebrates there was a negative relationship between the spatial distance between sites and the community similarity for species with low and moderate dispersal ability. This could mean that as the distance between suitable sites increases as more streams become environmentally degraded over time, the chances of such species re-colonising isolated streams will decrease. As discussed earlier, freshwater ecosystems are by their very nature prone to fluctuations in their stability; floods, droughts and other temporary stressors may lead to the local extinction of one or many of the local species present in the ecosystem. Without emigration from neighbouring streams within the region, it is hard to imagine a mechanism for the natural re-introduction of such species.

The composition of the community found in any given ecosystem is subject to the ability of individuals to reach that area. Neutral theory (Hubbell, 2001) suggests that

stochastic forces and dispersal limitation were more responsible for community composition than the niche based processes advocated by many community ecologists. While there is strong debate between opponents and proponents of the theory, Townsend and Townsend (2006) contend that it is more likely a combination of the two theories that best describes the processes that actually account for the make-up of real world communities. Their study of stream communities in New Zealand suggests that the wide range of dispersal abilities may play a vital role in determining the species present in any particular stream. By extension, in landscapes that are increasingly fragmented over time, ecosystem services provided by certain taxa may be absent, resulting in increased stress on the ecosystem.

Inbreeding is most often the result of small population size, and can result in reduced genetic diversity or inbreeding depression (Charlesworth and Charlesworth, 1987). In combination with the range of stochastic forces that may result in an increased stress level in the environment, inbreeding depression can increase a population's susceptibility to extinction (Reed, et al., 2002). It has long been thought that the effect of inbreeding and the resulting loss of heterozygosity could cause the extinction of a population, though for a long time there was little direct evidence (Saccheri et al., 1998). In a study on the Glanville Fritillary Butterfly (*Melitaea cinxia*), Saccheri et al. (1998) found that in this large meta-community as heterozygosity deceased, the likelihood of the local extinction of that patch would increase significantly. This was found to be especially true in small isolated communities, demonstrating the importance of regular emigration to the long term prospects of any given population. Dispersers emigrating from adjacent populations provide fresh genetic material that increases heterozygosity and by doing this increases the population's resilience against a myriad of other stressors.

The life history and dispersal of the mayfly *Coloburiscus humeralis*.

Freshwater insects spend the majority of their lives as larva or nymphs living within streams and rivers. Streams and rivers constitute discrete habitats embedded within a terrestrial landscape that is inhospitable to the juvenile stages, and for some taxa all life stages (Bilton, 2001). For mayflies, juveniles are restricted to freshwaters where they hatched because at this stage in their life-cycle they are a flightless. Harding and Winterbourn (1993) found that under some conditions the nymphs of *Coloburiscus humeralis* (Family: Coloburiscidae) could spent up to 27 months in a stream before emergence. Mayflies are unique among insects in having two winged life history stages, subimago and imago (Brittain, 1982). Nymphs emerge as a winged subimago and to avoid predation, quickly find a place in the terrestrial environment to moult and emerge as a final stage imago. Adult mayflies rely completely on fat reserves built up during the larval stage, and have lost functional mouthparts over evolutionary time (Brittain 1982). Because of this, the adult mayfly's life-stage is very short, up to two weeks, but for most species the winged stage lasts for as little as 24 hours (Brittain, 1989). In most mayfly species adults form large mating swarms, with males jostling for position within the swarm (Harker, 1992). Females fly into these swarms and mating occurs in flight. In most species swarming occur above some type of landscape feature, called a swarm marker. Wisely (1965) observed that swarms of *Coloburiscus humeralis* males formed over the stream in early morning or evening, or when there was a drop in wind speed.

There have been previous studies regarding the dispersal abilities of New Zealand mayfly species, and *Coloburiscus humeralis* has been a focus of at least two studies (Hogg et al, 2002; Morris, 2005). However both of these studies used a genetic marker, allozymes, that has been shown to be limited in its ability to resolve recent

dispersal. Both found that *Coloburiscus humeralis* had very low levels of genetic differentiation between populations (measured as F_{ST}) and therefore concluded that dispersal over long distances was a common occurrence for this species. However, allozymes are based on differences in the composition of amino acids within various proteins detected using an electrophoretic technique (Arnauld-Haond et al., 2005). Because the majority of DNA mutations do not alter protein composition very little variation can be expected, and therefore modern genetic structure is not easily resolved. Thus, dispersal in *Coloburiscus humeralis* is an area that warrants exploration using markers with a better ability to resolve modern patterns of gene flow.

Studying dispersal in freshwater insects

The dispersal of aquatic insects has been studied both directly, using passive and active traps (Didham et al., 2012; Kovats, et al, 1996; Jackson and Resh, 1989; Collier and Smith., 1995; Peterson et al, 2004.) and mark and recapture experiments (Caudill, 2003; McNeale et al., 2004; Briers et al., 2004; Svennson, 1974), and indirectly using molecular markers (Hughes, et al 2003; Hogg et al 2002). Overall, direct methods like trapping have shown that aquatic insects are most likely to be trapped close to their natal stream, with the number of adults collected in traps decreasing precipitously as distance from the stream increases (Collier and Smith, 1997), while the conclusions from genetic studies have been much more diverse (Alexander, 2011, Hogg et al 2002, etc).

Direct measurements of dispersal

Trapping

Prior to the widespread use of molecular tools to investigate dispersal, the primary methods for studying insect movement were the use of malaise and light traps. Using malaise traps, Petersen et al (2004) collected 29,000 caddisflies, stoneflies and mayflies and found that 90% of animals were trapped within 60 metres of a stream, they found that this was the case regardless of land use. The use of malaise traps in New Zealand has shown that mayflies and stoneflies were more likely to be collected within forested areas very close to streams, while caddisflies were found in a wider range of landscapes (Winterbourn et al., 2007). Long distance dispersal may occur most often due to random wind assistance (Milner et al 2000). Because of this, the vertical height that adults are found at might significantly affect dispersal distances. The vertical distribution of adult aquatic insects has also been studied using malaise traps. Collier and Smith (1995), set up malaise traps at three heights, 0.5m, 1.5m and 5m above the ground, and found that while caddisflies were caught consistently across all heights; mayflies were predominately caught at the 1.5m level. Furthermore, in the US male adults of the caddisfly, *Gumaga nigricula,* were found more regularly at traps that were set higher, while there was no relationship for females of the species (Jackson and Resh, 1989). Traps have also been used to assess the impact of road culverts, on the upstream movement of aquatic insects in urban watersheds. Blakely et al., (2006) found that the diversity and abundance of caddisflies was greater downstream than upstream, suggesting barriers to upstream dispersal.

Mark and Recapture

Experiments using mark and recapture have also been used to study the distance aquatic insect adults are able to fly. Svensson (1974) marked the forewings with permaclips and then released three species of caddisflies both close to and away from the stream and was able to recover individuals over 1 kilometre away from the release point. Another study, using the stable isotope ^{15}N as a label, tagged all the mayflies in a beaver pond and found that there was frequent dispersal between ponds, indicating that dispersal played an important role in the source-sink dynamics of the larger metapopulation (Caudill, 2003). In the stonefly *Leuctra ferruginea,* it was found that 76% of labelled adults were found upstream of the site where they were labelled, demonstrating that for this species adults predominately disperse upstream (Macneale et al., 2004). Briers et al., (2004) used the ^{15}N labelling technique to tag more than 1.5 million stonefly larvae was able to find direct evidence of this species (*Leuctra inermis*) dispersing from one river system to another. While this study showed that the distribution of recaptured adults coincided with prevailing wind patterns, suggesting an important role for wind assisted dispersal, the low numbers of recaptured adults limits the power of this study.

Limitations of Direct methods

Trapping studies have a number of shortcomings and rely on a number of assumptions. Firstly, the amount of fieldwork and large number of hours necessary to collect enough insects to make valid conclusions is costly and potentially prohibitive. Malaise trap studies are only able to take a snapshot in time, as the presence and abundance of an adult aquatic stream insect species is likely to

fluctuate through time, due to stochastic environmental variables. Furthermore, there is no guarantee that species caught in a trap have come from the stream being trapped, they may have come from another habitat further afield. Therefore, to make conclusions regarding underlying dispersal mechanisms and barriers, multi-year studies are needed. Furthermore, trapping studies require an abundant population of insects, with the likelihood that an insect will encounter a trap decreasing with population size (Collier and Smith, 1995). As was demonstrated by the Briers et al. (2004) mark and recapture study, a very large number of individuals is needed to be tagged in order to recapture even a very small proportion of them, severely limiting the conclusions that can be made.

Indirect methods for studying dispersal

Genetic Markers

In contrast to direct methods of dispersal, genetic markers provide information on the amount of gene flow occurring across a population. Gene flow can be defined as all of the mechanisms involved in moving of genes from one population to another (Slatkin, 1985). For mayflies, the primary means of moving genes from one spatially segregated population to another is via aerial adult dispersal. We can thus use genetic markers to infer levels of dispersal between such populations.

When choosing a genetic marker for a particular study, it is important to choose the one that best fits the aims, time-frames and budgets of that study. For studies that aim to quantify the levels of dispersal occurring in contemporary landscapes, it is

important to choose a marker that has a high mutation rate, so that the high mutation rate provides lots of variation and that drift and selection act on this variation (along with mutation creating new alleles) resulting in differentiation between populations. Many early population genetics studies on freshwater insects in New Zealand used the protein based allozyme marker technique. Because proteins have real phenotypic consequences on an organism, they are highly conserved, with low variation and a slow mutation rate. Allozymes also have low variation because many mutations do not change the amino acid composition of the protein. Therefore, markers such as allozymes are more suited to revealing ancient dispersal barriers and for use in phylogenetic studies rather than studies aiming to uncover contemporary barriers to dispersal (Bossart and Prowell, 1998). In a study of the caddisfly *Orthopsyche fimbriata,* Collier and Smith (2001) found no population level differentiation between populations separated by 10km but at the 100km range they found considerable differentiation. While there may be significant modern gene flow between the populations 10 kms apart, the slow pace of mutation of the marker used in this study may potentially be obscuring more modern barriers to dispersal between these populations. The two previous studies done on the population genetics of *Coloburiscus humeralis* (Hogg, et al., 2002, Morris, 2004) may have found very little differentiation within regions in Banks Peninsula in the South Island using allozyme markers. While these findings suggest that historically, *Coloburiscus humeralis may have* regularly dispersed between river catchments, they provide little information on contemporary dispersal patterns. In contrast, many studies using more rapidly mutating markers have shown that mayflies are often poor dispersers and can be quite strongly differentiated between populations at local scales (Gibbs et al, 1998; Stutz et al, 2010; Drotz et al, 2012).

Microsatellites are commonly used genetic markers that have a high enough mutation rate to be useful for contemporary studies of gene flow. Microsatellites are co-dominant markers and studies that use them usually select a few (5-20) highly informative multi-allelic loci (Meudt and Clarke, 2007). The co-dominant nature of these markers enables population geneticists to visualise both of the possible alleles at a particular locus in diploid organisms, meaning a precise analysis of heterozygosity and therefore estimations of inbreeding are possible (Sunnucks, 2000). However microsatellites do have some drawbacks; primarily that primer development for microsatellites can be a costly and time consuming exercise because of the need for sequence data to inform primer generation (Guichoux et al., 2011). In contrast, Amplified Fragment Length Polymorphisms (AFLPs) are able to generate highly informative polymorphic datasets with a much lower threshold to entry in terms of both time and monetary cost.

Amplified Fragment Length Polymorphisms (AFLPs)

Originally developed by Vos et al., (1995), AFLP PCR is a genetic technique that is able to detect various polymorphisms in different genomic regions simultaneously. Because of its ability to amplify a large number of fragments at one time, its reported sensitivity and reliability, it has been used in a variety of studies. Additionally, because the AFLP process creates a genomic survey of these diverse genetic fragments using a restriction-ligation technique, no *a priori* sequence data is necessary. This means that genetic studies can be carried out on non-model organisms relatively inexpensively (Meudt and Clarke, 2007). Because of the high number of fragments assessed in any given AFLP based study, with these fragments

coming from across the genome, there is a high likelihood that many will originate from non-coding areas of the genome and therefore are often quite variable (Shirasawa, et al. 2004). Because of this tendency towards highly variable genotypes, AFLPs lend themselves best to studies of recent population divergence and shallow phylogenies rather than questions relating to deeper evolutionary time (Meudt and Clark, 2007).

The AFLP technique creates dominant datasets; that is the alleles at any particular locus can have one of two states, present (1) or absent (0). A present AFLP peak means that at least one copy of the locus amplifies in PCR, while an absent peak means that both copies do not amplify. Therefore, with dominant markers such as AFLPs we are unable to distinguish individuals that are present for both copies (homozygote) from individuals present for one copy only (heterozygotes). While co-dominant markers such as microsatellites are able to give us direct measures of heterozygosity because of their ability to measure the allelic makeup of each locus, dominant data is less informative and requires many more loci to make conclusions regarding changes to allele frequencies between populations, and thus dispersal. This also means that, without an accurate measure of heterozygosity, estimating levels of inbreeding is difficult with dominant markers.

Assumptions of AFLPs

Like all genetic markers, interpreting AFLPs requires a number of assumptions. Firstly, because of the anonymous nature of the fragments that the AFLP technique produces, we must assume that all fragments that are the same length have an

identical nucleotide sequence. It has been shown that some primer combinations may produce higher levels of size homoplasy (two or more fragments of the same length but different genomic origin are scored as the same allele) than others, with the probability of finding size homoplasy as the number of fragments per primer set increases. The result of increased size homoplasy in a data set can be an underestimation of the differentiation between sub-populations, and therefore a decrease in the power of the study to detect the genetic variation in the wider population (Caballero et al., 2008). Therefore, it is important to be aware that AFLP data sets that show little genetic differentiation may not necessarily mean that there are no barriers to dispersal, and in such cases further investigation of the levels of size homoplasy may be required (Caballero et al., 2008).

Secondly, AFLPs are assumed to be selectively neutral. Because we are trying to measure the amount of genetic differentiation between populations as a result of a dispersal barrier, there should be no selection acting on the AFLP markers that favours one allele over another, potentially masking patterns of gene flow. When comparing a randomly bred group of catfish to groups selected by body weight Mickett et al. (2003) found almost identical reduction of AFLP polymorphism as population numbers declined, suggesting that AFLPs are selectively neutral markers.

Lastly, various programs that have been developed to analyze genetic data, such as Structure (Pritchard, 2004) make the assumption that the population is in Hardy-Weinberg equilibrium. Krauss (2009) suggests that departures from Hardy-Weinberg equilibrium may be modest in dominant markers such as AFLPs in out-crossing species. Hardy-Weinberg equilibrium makes the following assumptions (Klug and Cummings, 2003);

1. There is no selection amongst individuals of all genotypes.
2. There are no mutations occurring.
3. The population is infinitely large.
4. All mating is random.
5. There is no migration in or out of the population.

While many of these assumptions are unrealistic in natural populations, they can be used as a null hypothesis to compare to a genetic dataset. While assignment testing using the allele frequency based algorithm in the program STRUCTURE 2.2 requires the assumption of Hardy-Weinberg equilibrium, Falush et al, (2007) found that in a number of dominant datasets that these assumptions were met due to a new statistical technique included in the latest version of the software.

Landscape genetics

There has been an increased emphasis upon connecting patterns found in population genetic studies with the landscape in which the populations are imbedded to identify dispersal barriers, find dispersal corridors and determine the impact of anthropogenic changes to the landscape on the ability of organisms to successfully disperse. Statistics of genetic distance such as F_{st} and its analogue ϕ_{st} are able to give us information on the amount of genetic differentiation among populations, but can tell us little about which geographic and ecological processes are generating this genetic structure. To address this deficiency, landscape genetics integrates a variety of disciplines, including ecology, geography, spatial statistics and population genetics, to investigate the links between the landscape and genetic variation of populations (Manel et al., 2003; Storfer et al., 2010.)

To examine how landscape variables affect gene-flow and spatial structure in heterogeneous environments, a number of different modeling techniques have been employed by a rapidly growing number of recent studies. Least-cost analyses have been widely used, but can be restrictive as they assume dispersing individuals choose a single optimal path (McRae et al., 2008). More recently, McRae and Beier (2007) employed a resistance modeling technique based on circuit theory, which has been since used in a variety of papers to successfully answer a broad array of questions. This includes: finding specific barriers to gene-flow (Shwartz et al., 2009), finding connective pathways through landscapes (Shafer et al, 2012), and identifying the effect of anthropogenic changes on populations of threatened species (Sackett et al, 2011).

Circuit Modeling

Resistance modeling uses electrical circuit theory to create resistance surfaces that assigns different resistance values to various landscape features. If a particular cell in the rasterized landscape grid is assigned a high resistance value, then the amount of opposition the landscape feature presents to the movement of the organism will be high. Conversely, cells that are given low resistance values will have high conductance, in ecological terms this means that these areas will be pathways for movement across the landscape (McRae et al., 2008). Resistance modeling has a solid basis in random walks theory (Doyle and Snell, 1984), and unlike least-cost models is able to detect multiple connective pathways through the landscape, and therefore provides the opportunity to discover novel dispersal corridors that the researcher had not considered prior to the study.

In this study I used a multi-model factorial approach similar to the one developed by Cushman et al., (2006) to evaluate a number of alternative hypotheses and to identify the combination of landscape factors that most affect the gene-flow of *C. humeralis* across the Arthur's Pass study area. Because it is questionable to make assumptions regarding the relative importance of factors *a prioiri,* an additive multifactor model is the best course of action for resistance modeling (Cushman et al., 2006). In this way, I was able to not only evaluate the potential effect of gene-flow caused by a single factor, but by adding them together in all potential combinations I will be able to evaluate how these landscape variables may interact to create the pattern that is the closest fit for any genetic structure found among my sample locations. By using an additive framework, we can address potential problems of correlation between factors. For example, ridgeline barriers and open areas may be correlated, as ridgelines are by definition open areas. However using an additive model we can test if the ridgeline is a barrier above and beyond the impact of open areas, thus adding the elevation/slope element into the model.

Research Aims

The objective of this research project was to determine the dispersal patterns of the endemic mayfly species, *Coloburiscus humeralis*. Specifically:

- Are there barriers to dispersal between sub-populations living in isolated streams?
- If so, what specific components of the landscape are creating any such dispersal barriers?

- To what extent is dispersal of *Coloburiscus humeralis* affected by lack of suitable habitat between streams?

The rationale for choosing this species was twofold; firstly it is a widely distributed and when present, locally abundant species, found throughout New Zealand from Northland to Southland (Hogg et al, 2002). This may indicate that dispersal between catchments for this species was historically frequent and widespread; as dispersal limitation over large timescales in a pre-anthropogenic age would like have led to local adaptation and likely various speciation events (Hogg et al., 2002).

I hypothesised that as was found for the North American mayfly, *Ephemeralla invaria,* (Alexander et al., 2011.) decreasing forest connectance between streams would lead to increased genetic differentiation, because of the potential barrier to dispersal that contemporary deforestation may provide for these weak flying insects. To do this I collected insects from streams within Arthur's Pass National Park, an area with high percentage of forested area connecting streams. I aimed to compare the genetic structure of these streams with streams from area outside of the park that is characterised by forested patches surrounded by open high country farmland. Then, using the landscape genetics technique of circuit modelling (McRae et al., 2008), I sought to provide a quantitative measure of the differences in gene flow due to various landscape features, including forest cover.

Chapter 2
Methods

Experimental Design and stream sites

Individuals of the mayfly species, *Coloburiscus humeralis* (Family Coloburiscidae) were collected from 10 streams from across New Zealand; 9 in the South Island and one in the North Island. At the largest scale, the study streams were divided into three geographically distinct regions (Arthur's Pass, Banks Peninsula, and Northland). The main study area, the Arthur's Pass (AP) region located in the Southern Alps of the central South Island contained seven sites, while two sites were located on Banks Peninsula (BP) and an additional stream was sampled in Northland (Nl) (Figures. 2.1, 2.2, 2.3). The Arthur's Pass populations were central to my analysis of population structure and dispersal patterns at the local scale, while the Banks Peninsula sites were added to confirm that genetic differentiation was also occurring between populations in a separate region. Additionally, the Banks Peninsula sites were used to investigate the amount of gene-flow occurring between these populations and the Arthur's Pass populations. The two regions were separated by 120 km, with the intervening agricultural landscape of the Canterbury Plains, which has limited mayfly habitat. Lastly, the Northland site was 780 km away and on another island, and was included in the analysis as an out-group population, with the expectation there should be a significant amount of genetic differentiation between this population and the others.

The Arthur's Pass region includes open pasture and forest which is dominated by beech (*Nothofagus* spp.), while open high country farming occurs outside Arthur's Pass National Park. Three streams were sampled within Arthur's Pass National Park. The high country beech forest of the National Park is separated by montane areas above the bush line, and the open riverbeds of the Hawden, Poulter and Bealey rivers. Three streams in continuous forest, Peacock (Pa), Moss (M), and one unnamed stream that I have called Waimak Bend (Wb), had populations of *Coloburiscus humeralis*. While these streams are separated by 9.59 kms (Table 2.2), I sampled over 50 first and second order streams in the National Park that were connected by continuous forest and not separated by road, and it was only these streams that I found the nymphs of *C. humeralis* (See Figure 3.1).

Three streams (Castle Hill, Craigieburn and Pylon Gully) that were sampled ran through fragments of beech forest and one stream, Cheeseman stream (Ch), was characterised by riparian regenerating scrub consisting of bracken (*Pteridium esculentum*), exotic grasses and some scattered mānuka (*Leptospernum scoparium*). All of these streams were separated by intervening high country pasture (See Figure 2.1).

Banks Peninsula streams were located within regenerating stands of mixed podocarp forest. While the Northland stream (Pukenui forest near Whangarei) is predominately podocarps, with species such as Kahikatea (*Dacrycarpus dacrydioides)*, Totara (*Podocarpus totara*) and the Laurel, Taraire (*Beilschmiedia tarairi)* occupying the canopy.

Insect Collection

Mayfly nymphs were collected in the summers of 2011/2012 and 2012/2013 and in April and May, 2013. At each site *Coloburiscus humeralis* nymphs were collected from a 25m reach of the stream using a kick-net (250μm mesh) and hand-picked from substrate. Nymphs were then preserved in vials of 95% ethanol and returned to the laboratory, where they were stored in the freezer at -18°C.

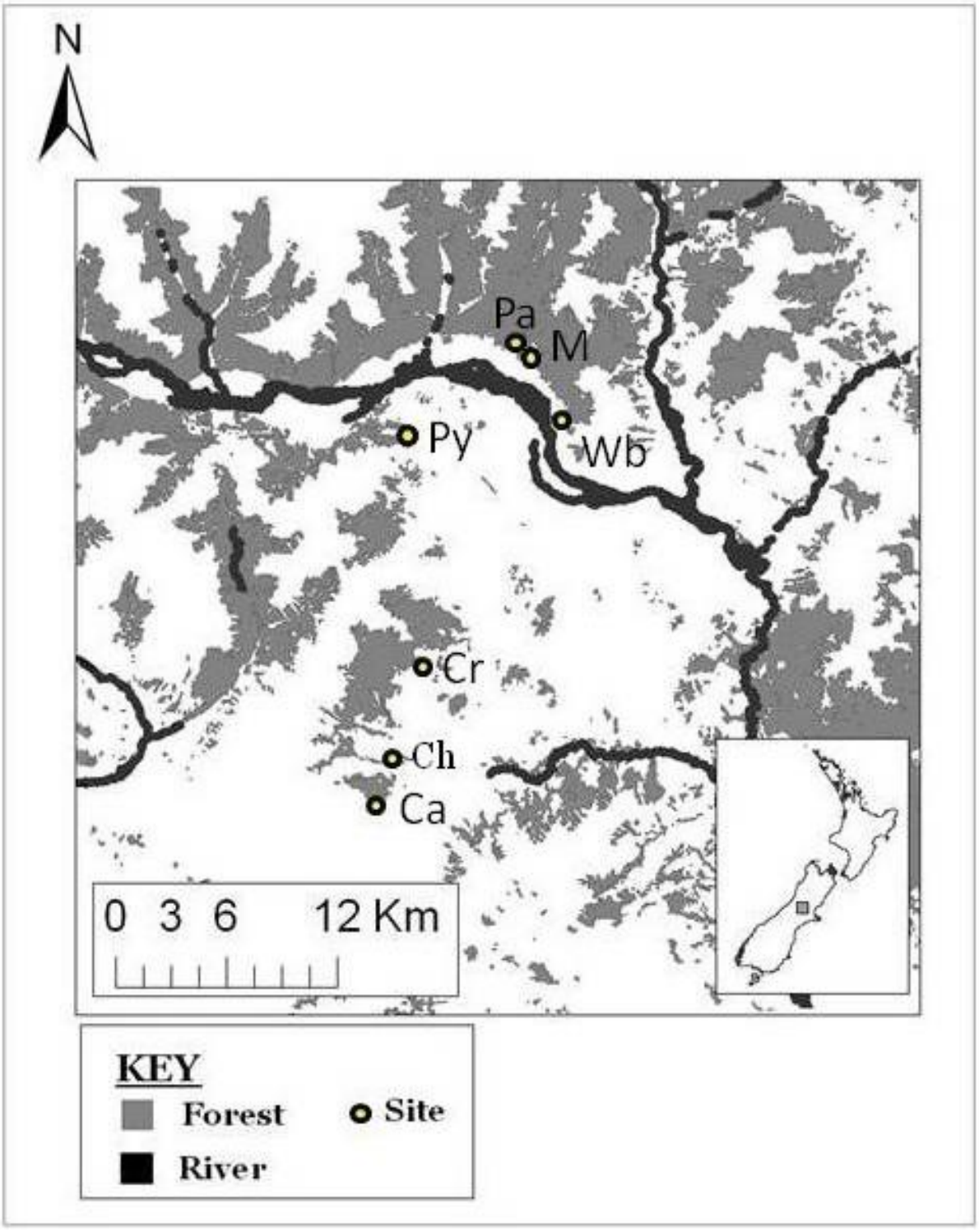

Figure 2.1 Map showing the location of the seven Arthur's Pass sampling sites. Ca = Castle Hill, Ch=Cheeseman, Cr = Craigieburn, Py = Pylon Gully, Pa = Peacock, M = Moss, Wb = Waimak bend.

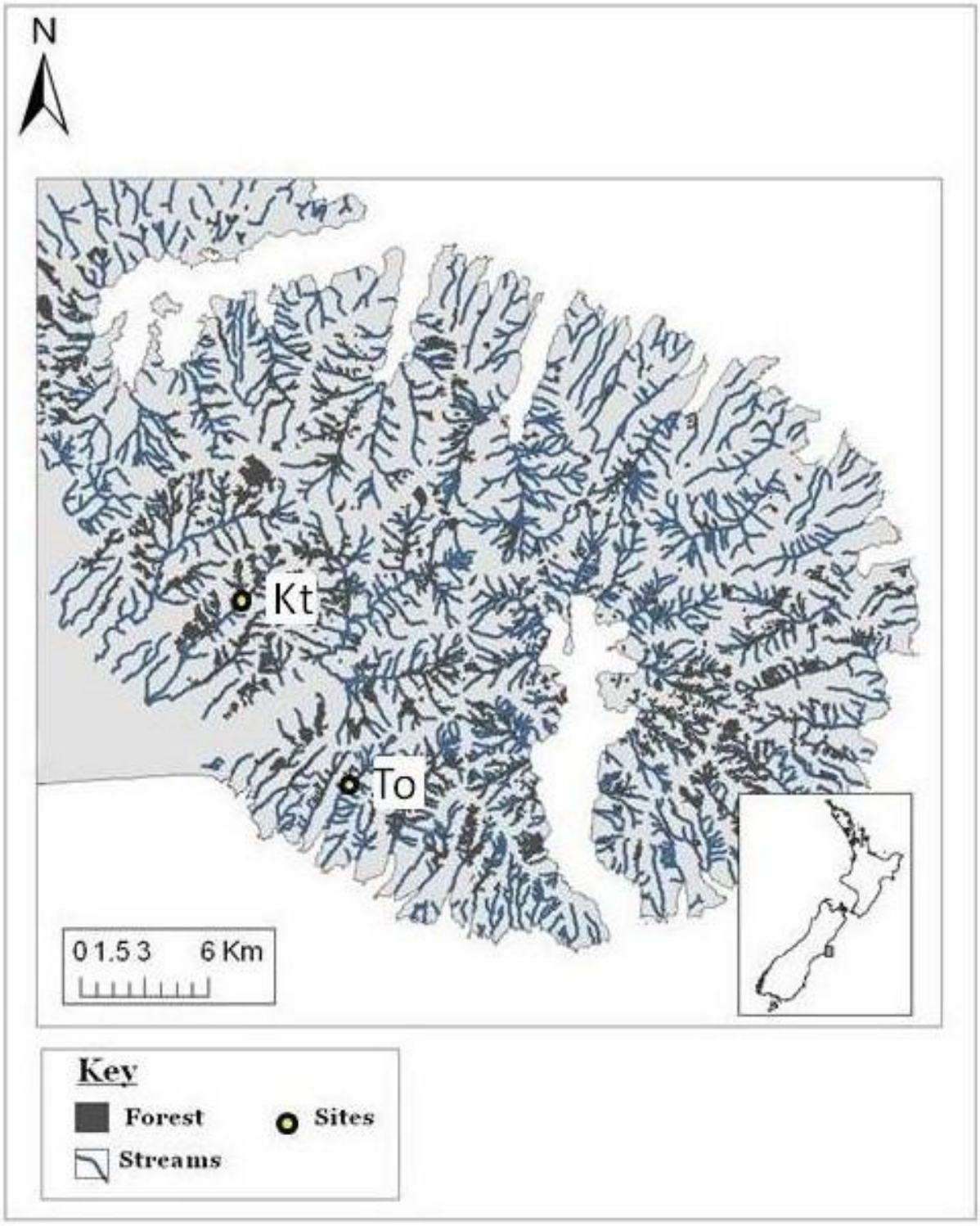

Figure 2.2 Map showing the location of the two Banks Peninsula sampling sites. Kt = Kaituna, To = Te Oka.

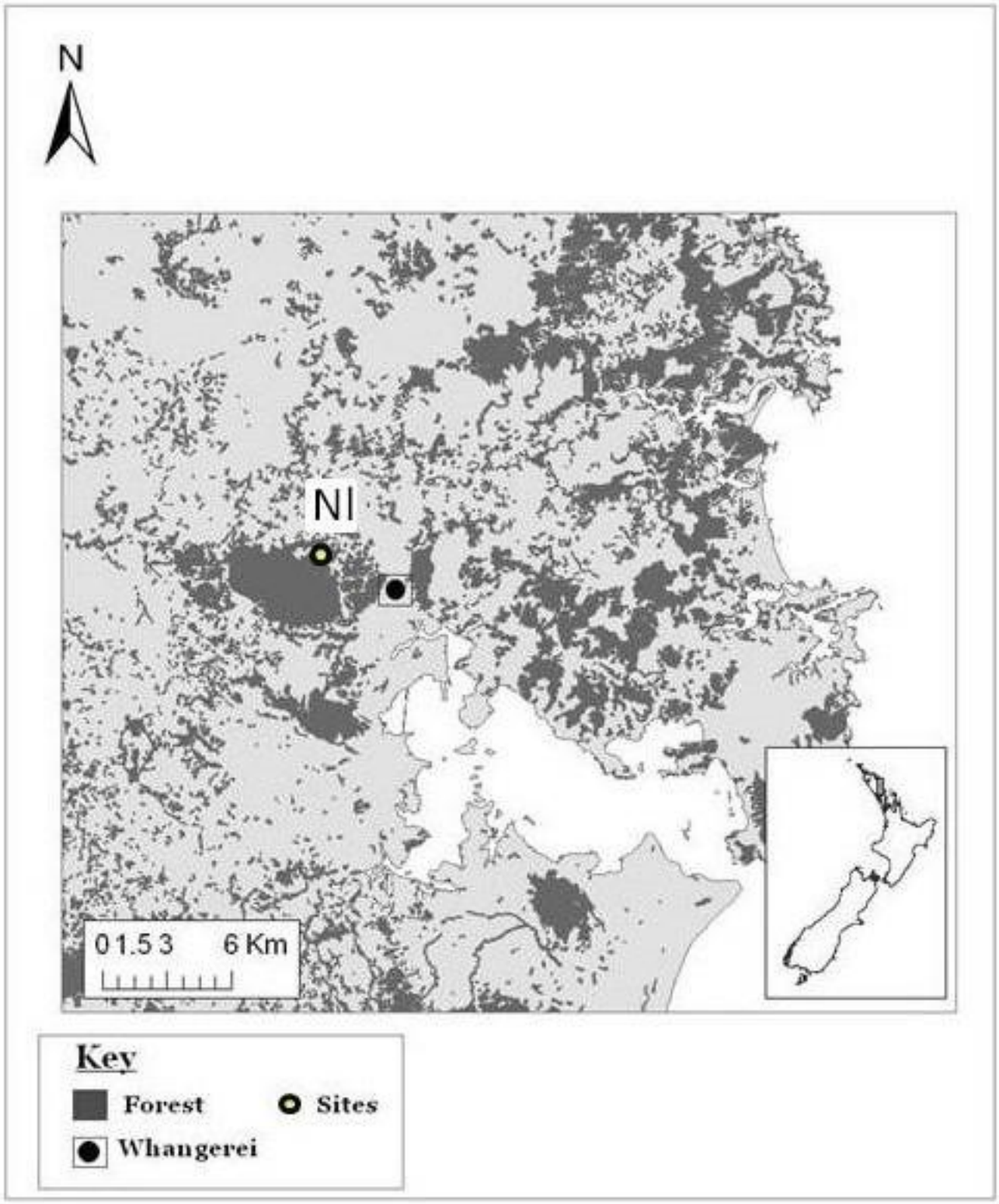

Figure 2.3 Map showing the location of the sampling site in Northland.

Nl = Northland.

TABLE 2.1 : Pair-wise Euclidean Distance (Kilometres). Using the geographic distance calculator in GenAlX (Peakall and Smouse, 2012). AP= Arthur's Pass. BP= Banks Peninsula. NL= Northland

Sites	**Region**	**Castle Hill**	**Cheeseman**	**Craigieburn**	**Pylon Gully**	**Peacock**	**Moss**	**Waimak bend**	**Kaituna**	**Te Oka Bay**	**Northland**
Castle Hill (Ca)	AP	0									
Cheeseman (Ch)	AP	2.6	0								
Craigieburn (Cr)	AP	6.7	4.2	0							
Pylon Gully (Py)	AP	17.6	15.3	11.3	0						
Peacock (Pa)	AP	23.7	21.2	17	8.2	0					
Moss (M)	AP	23.4	20.9	16.7	8.7	1.2	0				
Waimak bend (Wb)	AP	22.3	19.7	15.7	9.6	3.8	2.6	0			
Kaituna (Kt)	BP	114	113.7	114.5	121	118.2	116.9	114.4	0		
Te Oka Bay (To)	BP	125.5	125.4	126.4	133.5	131	129.8	127.2	14.1	0	
Northland (Nl)	NL	793.4	790.9	786.7	776.2	769.7	770.1	771.6	816.5	828	0

AFLP Analysis

DNA Extraction

I initially employed a CTAB protocol (Weising et al, 1995) to extract DNA from the collected *C.humeralis* larvae. The CTAB method is popular in population genetic studies because it is able to produce large quantities of DNA, while using typical laboratory chemicals and therefore remaining inexpensive (Reineke, 1998). However, it has been noted that DNA extracts obtained using the original CTAB protocol may contain high levels of contaminants such as polysaccharides, RNA, and phenols (Calderon-Cortes et al, 2010). The successful generation of AFLP genotypes requires both high levels of DNA and low levels of contaminants (Meudt and Clarke, 2006). To address this issue, Calderon-Cortes et al., (2010) suggest including PVP and β-mercaptoethanol, resulting in DNA extracts that had far lower levels of these contaminants when these two reagents were added to the extraction process. However, despite following these corrective steps, I later had trouble consistently generating informative and clear AFLP profiles using these DNA extraction techniques.

Commercial kits such as the PureLink Invitrogen Kit (Invitrogen Ltd) are often more expensive than CTAB based extraction techniques, and because this technique is based on a filter based elution method often produces smaller quantities of DNA, because not all target DNA passes through the filter. However, sufficient levels of DNA were generated using this kit (80ng/mg - 300ng/mg) and using this extraction kit I was able to more consistently produce reliable AFLP genotypes.

PureLink Protocol

In the laboratory, 5mm² of mayfly tissue from each collected individual was put in a 1.6 ml Eppendorf tube containing 200µl of digestion buffer. Each individual was then ground thoroughly with a freshly cleaned micro-pestle. 20µl of Proteinase K was added to the tube in order to help in the breakdown of potentially problematic proteins and was then vortexed well. The tubes were then left in a heat bath over-night at 55°C to allow the enzyme time at the optimum temperature to breakdown as much of the protein as possible.

After a period of at least 12 hours, 20 µl of RNASE A (an enzyme that breaks down RNA) was added to each of the sample tubes and then mixed thoroughly with a vortex. This product was then added to a spin column and spun in a centrifuge for 1 minute at 10000 rpm (revolutions per minute). The filter at the top of the spin column then contained the target DNA and any remaining contaminants. The filter was then added to a new collection tube and spun for an additional 1 minute at 10000 rpm. In order to remove any remaining contaminants that may be present two wash buffers were sequentially put through the spin column, with an intervening spin step at 10000 rpm for 1 minute. After the second wash buffer was added the column was spun for 3 minutes at the maximum speed of the centrifuge (14000 rpm). The spin column was then added to a sterile Eppendorf tube, and 50 µl of Elution buffer was added to the top of the column. During the final spin (1 minute at 10000rpm), the low salt Elution buffer dislodges the DNA that has been bound to the filter and the liquid that remains in the Eppendorf collection tube contains purified DNA. All of this was carried out in concordance with the instructions of the manufacturer (lifetechnologies.com).

Amplified Fragment Length Polymorphisms (AFLPs)

The following AFLP protocol is based on the work of Vos et al (1995).

Restriction

AFLPs are generated by the complete digestion of nuclear DNA using two restriction enzymes, a frequent and an infrequent cutter. Initially, the enzyme PstI was used in combination with EcoRI to restrict the genomic DNA. However, there was some difficulty getting this combination to completely restrict the genomic insect DNA. Therefore, the PstI enzyme was exchanged for the more expensive MseI restriction enzyme, and immediately was able to produce AFLP profiles as expected. MseI is the frequent cutter and cuts at the 4 base pair (bp) consensus sequence of AATT, while EcoRI cuts at the 6 bp consensus sequence of GAATTC. To run the restriction digest, I added ~0.4µg of template DNA from each sample, 5µl of the NEB EcoRI reaction buffer [100 mM Tris-HCl, 50 mM NaCl, 10 mM $MgCl_2$, 0.025% Triton® X-100], 5 units of each of the enzymes MseI and EcoRI and 9µl of PCR grade water to standard PCR tubes. When using two restriction enzymes in the same reaction, there is the potential that without the optimum buffer, the two enzymes may interfere with each other to impair the optimal restriction of the DNA. To avoid this I used the EcoRI buffer, as recommended by the manufacturer when these two enzymes are used in combination (ref-Invitrogen.com). Tubes were then mixed well on a vortex, before they were put in a thermocycler at 37°C for 4 hours. After this, an additional 15 minute cycle at 72°C was included to completely inactivate the enzymes to prevent any interference in subsequent reactions (Vos et al., 1995).

Ligation

The second step in the process is ligation, in which the enzyme ligase is used to attach complementary sticky ended pieces of DNA to each end of the fragment. This is possible because we know at what sequence the restriction enzymes cuts the genomic DNA, and therefore are able to design complementary double stranded adapters to attach to the end of each fragment. As a result, these fragments have at each end a known sequence of DNA that primers can be designed to attach to in the following PCR reactions.

The oligonucleotides, 5' GAC GAT GAG TCC TGA G 3' and 5' TAC TCA GGA CTC AT 3' were combined to create the MseI end double stranded adapter and 5'CTC GTA GAC TGC GTA CC3' and 5'AAT TGG TAC GCA GTC TAC3' were combined to create the EcoRI double stranded adapter. Firstly, I suspended each of the single stranded adapters to 100µM using TE buffer. A second dilution was done using PCR grade water so that each adapter solution was 10µM. I then combined 50µl of each of these single strand adapter solutions in a 1.7mL Eppendorf tube, mixed them well, and left them at room temperature for an hour to allow time for the adapters to anneal. 20 µl of a solution containing 5 pM EcoRI-adapters, 50 pM Msel adapters, 1 unit of T4 DNA-ligase, 1 mM of ATP in 10 mM Tris-HAc pH 7.5, 10 mM MgAc, 50 mM KAc, 5 mM DTT, 50 ng/µl BSA and ~400ng of restriction product was put in a thermocycler overnight at 37°C (Vos et al., 1995).

Pre Selective PCR

In each tube, PCR buffer [100 mM Tris-HCl (pH 8.3), 500 mM KCl], 2.0 mM MgCl2, 0.2 mM dNTP, 0.4 units of *Taq DNA* polymerase, 10 µM EcoRI+A primer, 10 µM MseI+C primer and 5µl of ligated DNA product were added, making the total reaction volume up to 20µl with PCR grade water. The tubes were then mixed thoroughly and put on the thermocycler with the following PCR cycle; 72°C for 2 min followed by 20 cycles of 94°C for 20 sec, 56°C for 30 sec, and 72°C for 2 min, and a final incubation of 72°C for 2 min and 60°C for 30 min.

Selective PCR

A second PCR was run to amplify a subset of the fragments generated in the pre-amplification step. Selective primers had an additional two nucleotides added to the sequence used for pre-amplification. The selective EcoRI primer was fluorescently labelled with 6-FAM (Invitrogen Ltd) to enable visualisation in the subsequent genotyping step. 8 MseI selective primers were trialled in combination with the one tagged EcoRI primer, in order to find the combinations that produced the most consistent and reliable genotype profile (Table 2.2).

1X PCR buffer [100 mM Tris-HCl (pH 8.3), 500 mM KCl], 2.0 mM MgCl2, 0.2 mM dNTPs, 0.625 µM of 6-Fam dye labeled EcoRI primer (E+3), 0.625 µM MseI primer (M+3), 0.4 units of *Taq* DNA polymerase, 2 µl of diluted preselective amplification product and PCR grade water was added to make a total reaction volume of 20µl used for selective amplification. The PCR amplification was carried out with an initial denaturation step of 94°C for 2 min, followed by the first cycle of 94°C for 20 sec, 66°C for 30 sec, 72°C for 2 minutes and one degree decrease in annealing temp in

each of the next nine cycles. This was followed by 25 cycles of 94°C for 30 sec, 56°C for 30 sec, and 72°C for 3 minutes (Vos et al., 1995).

Genotyping

AFLP bands were visualised using the Applied Biosystems Ltd **3130xl** Genetic Analyzer at the School of Biological Sciences, The University of Canterbury. In preparation for the genotyping, 12 µl of HiDi, 0.3µl of size standard (Life Technologies Ltd), and 1µl of selective PCR product were added to the genotyping plate. This was then run on the thermocycler for 3 minutes at 95°C in order to ensure complete denaturation of the fragments of DNA before genotyping.

Table 2.2: Primers and adapters used in AFLP genotyping. *Denotes adapters and primers used in successful genotyping.

Primer/adapter	**Tube name**	**5'-sequence-3'**	**Label**
EcoRI adapter 1	EA1	AATTGGTACGCAGTC*	
EcoRI adapter 2	EA2	CTCGTAGACTGCGTACC*	
MseI forward adapter	MA1	GACGATGAGTCCTGAG*	
MseI reverse adapter	MA2	TACTCAGGACTCAT*	
EcoRI selective primer labelled	ESP	GACTGCGTACCAATTCAG*	6-FAM
MseI selective primer 1	MSP1	GATGAGTCCTGAGTAACAG*	
MseI selective primer 2	MSP2	GATGAGTCCTGAGTAACTG	
MseI selective primer 4	MSP4	GATGAGTCCTGAGTAACAC	
MseI selective primer 5	MSP5	GATGAGTCCTGAGTAATAG	
MseI selective primer 6	MSP6	GATGAGTCCTGAGTAACCT	
MseI selective primer 7	MSP7	GATGAGTCCTGAGTAAGTC	
MseI selective primer 8	MSP8	GATGAGTCCTGAGTAACCC*	

Replication and scoring of AFLP Peaks

AFLP chromatograms were visualised using the software GENEMAPPER (Applied Biosystems Ltd). AFLP peaks were scored as either present (1) or absent (0), as they are dominant markers. 10 individuals, randomly selected from across the sampling sites were replicated 3 times each from the original DNA extraction. It was not possible to use independent extractions as the entire sample was used in the original extraction. This was done to ensure a high level of reliability that AFLP profiles remained the same across multiple laboratory runs and to estimate the error rate.

The scoring of AFLP profiles can be a somewhat subjective exercise, and it is therefore important that there is a simple and replicable scoring procedure that is used across all individuals to minimise the potential for errors that may obscure the actual genetic pattern in the population. Bonin et al. (2004), suggest that any peak that falls below 10% of the highest peaks intensity may not be reliable and should be excluded from the final analysis. However, when comparing the replicated profiles in this study I found that peaks with an intensity below 50 relative fluorescent units (rfu) were quite variable between replicates and because of this peaks below this level were excluded. Another area where the presence of peaks was quite variable between replicates was where small peaks were in close proximity to a larger peak, often appearing connected in the chromatogram. These shoulder peaks are considered artefacts of the genotyping process, and it is recommended that they be removed (Herrman et al, 2013). It is also possible that slight stutter can occur in the genotyping process, with peaks moving up to a base pair in either direction (Herrman et al, 2013). To counter this peaks that fell within a 1.5bp range of each other were binned together. While this is slightly conservative because it reduces the potential amount of variation, I decided it more important to ensure that if there was

a pattern of differentiation in my data-set it was a reliable one. Finally, I only scored fragments between 90bp and 500bp in length, as it has been noted that peaks in this range are the most reliable (Bonin et al 2007).

Genotyping failure and troubleshooting

After successfully generating AFLP profiles using 2 of the 10 selective primer combinations that were trialled (MseI+CCC, MseI+CAG; in combination with the fluorescently labelled EcoRI primer), there was an unexpected failure to produce any usable peaks in subsequent genotyping runs. All chromatograms that I analysed at this time were close to identical, including the previously clear negative controls. The negative controls were run through all of the steps to produce AFLP profiles with only the first restriction step not receiving any insect DNA. All profiles that were generated in this run had the same 2 very large peaks (>10000 rfu) at the 64 bp and the 83bp length. Because these two peaks occurred in the negative control as well as the others, there must have been some laboratory based contaminant that may have been inhibiting one or more of the reactions in the AFLP process. Furthermore, Bonin et al. (2004) suggest that contamination is a common factor in the failure of the AFLP process.

To test this possible contamination, the first step that I tried was to completely clean and then autoclave my pipettes. Because of the possibility of cross contamination between transferring DNA to wells and transferring AFLP reagents to master-mix tubes, from this point forward I used one pipette solely for DNA transfer and the rest for preparing and applying the master-mix to PCR wells. However, despite these

precautions, the negatives still showed the 2 very high peaks at 64bp and 83bp in length.

Bonin et al. (2004) recommend separating master-mix creation from the handling of DNA into two separate areas of the laboratory. For the creation of master-mixes for each reaction I used a positive pressure fume hood located in an adjacent laboratory. Before use I ensured that the hood was thoroughly cleaned with bleach and 70% alcohol. However, this precaution did not generate clean negative controls.

At this point I was reasonably confident that the instruments that I was using were sterile and I had reduced the risk of cross-contamination sufficiently. It therefore seemed likely that one or more of the reagents used in the AFLP process had been contaminated prior to the subsequent more rigorous cleaning measures. I therefore ordered fresh stock of each of the solutions in the process. While I was able to acquire fresh stock of all PCR reagents immediately, primers and enzymes took longer to arrive from the suppliers. After replacing the PCR reagents with fresh stock I was able to produce clean negatives, however despite this, AFLP fragments were not able to be visualised in from the genotyping run.

Therefore, it was apparent that one of the steps in the process was not working as it previously had. Because of the fresh aliquots of PCR primers and reagents that I had access to at this point, I was reasonably sure that it was not the PCR step that was causing the problem. I therefore investigated whether there was any part of the first three steps that was causing the genotyping failure.

At this stage my DNA extractions were several months old as I had extracted them as a batch in the very beginning of my laboratory work. Successful AFLP generation requires high quality DNA (Meudt and Clarke, 2006), and if there had been a

degradation of the extracted DNA there may not have been sufficient template for subsequent steps to cut, label and amplify. Using the spectrophotometer NANODROP (Thermo Scientific ltd), I assessed the quantity and quality of the original DNA extractions. I found that there was in almost every case no remaining DNA in the extraction tubes.

I then returned to the field to collect fresh samples, and extracted them using the common CTAB-PVP extraction technique (Rogers and Benditch, 1985). I then ran these samples through the spectrophotometer and found a good quantity and quality of DNA (100-400 ng/µl, 260/280 = ~2.00). To ensure that there were as little unwanted contaminants such as RNA in these extractions as possible, I performed a gel electrophoresis on these unamplified DNA extractions. I found that while there was a strong band at around the 8000 bp range on the agarose gel, there was often a smear at the small bp end of the gel. This smear is not a part of the target genome, and is likely a contaminant such as RNA, or another DNA source. I then bathed the new extractions in 20µl of RNAse at 37°C for 20 minutes, to attempt to breakdown any residual RNA in the tube. Gel electrophoresis of these samples no longer showed the smear that had been present previously. This would suggest that the smear on the electrophoresis gel had been RNA. This demonstrated that while I could be reasonably sure that RNA was not going to further inhibit the genotyping process, the CTAB extraction method may not remove other inhibiting compounds to the extent required for AFLP genotyping, especially compared with a filter based extraction kit. I then used the PureLink extraction kit (Life Technologies Ltd), which provided slightly lower DNA yield but showed no sign of contamination in either the spectrophotometer or when a gel electrophoresis was run.

To test whether the restriction enzymes were cutting as expected I tested them on the viral Lambda genome. Lambda DNA is model laboratory genome that is known to have cleavage sites for both MseI and EcoRI. I ran the restriction of the lambda DNA using the same protocol that has been outlined above, except that two restrictions were run simultaneously with one restriction enzyme for each reaction. This was done because a manufacturer (ThermoScientific Ltd) has a model agarose gel that visualises the restriction profile of EcoRI and the enzyme Tru1I (an analogue of MseI) on Lambda DNA. There are 6 cleavage sites on the Lambda genome, and 195 for MseI (www.thermoscientificbio.com). I was then able to produce agarose gels that matched the cleavage patterns as provided by the manufacturer, suggesting that the restriction enzymes are active and therefore should be able to cleave the insect DNA as they had done before. Additionally, I ran restrictions on the *C.humeralis* DNA and then visualised them using gel electrophoresis. While the laboratory based electrophoresis rigs are not able to resolve fragments of the length that these restriction enzymes are likely to generate, I was able to see that there was no longer a large strong band at the 8000bp region, and it was replaced in all the samples by a larger smear between the 100bp and 700 bp size range. This suggests that the restriction enzymes were digesting the insect DNA as expected.

Because the main purpose of the ligation step is to attach very short lengths of DNA to the end of each of the fragments, there was no way to use electrophoresis or any other at hand laboratory technique to assess whether it had been successful independently of the other steps. However, at this stage I was confident that all of the other steps should be working, and because I was still not generating recognisable AFLP profiles; I was confident that the main cause of this was that one of the ligation components that had degraded. Furthermore, it has been previously

reported that the ligation step has been known to be a source of genotyping failure for AFLPs, due to of the propensity of the ATP that is part of the solution to degrade (Treier, 2003).

After I received a new stock of ligase, I was able to produce 209 strong, replicable peaks for 161 individuals from all 10 of my study sites. However, while I was able to amplify DNA fragments using the MseI-CCC selective primer, I was unable to produce any genotypes using the MseI-CAG selective primer combined with the fluorescently tagged EcoRI selective primer after the initial genotyping failure. Furthermore, while I collected 30 individuals from each site I was unable to produce genotypes for all the animals that I collected. After genotyping the first few batches of samples (160 individuals), I again found that the profiles that I was producing did not carry sufficient peaks for accurate analysis. While I suspected that this was again due to fragility of the ligase, time and money constraints prevented me from genotyping any more of the samples.

Statistical Analysis

Measures of population heterozygosity and the percentage of polymorphic loci for each of the study sites that were sampled were calculated in AFLP-SURV (Vekemans et al, 2002). While the percentage of polymorphic loci (PLP) is a simple measure of the proportion of loci in each population that are polymorphic, deriving expected heterozygosity (Hj) from dominant data requires a few assumptions. Lynch and Milligan (1994) make a number of assumptions when estimating expected heterozygosity (Hj), an analogue of the (He) measure of heterozygosity used in co-dominant studies (Vekemans et al, 2002). Firstly, to estimate Hj, Lynch and Milligan (1994) make the assumption that marker alleles of different genetic origins

do not co-migrate to the same banding position, an assumption they call generous. It is important to note that this is a definite limitation of dominant markers, AFLPs clump all fragments of the same length together and are unable to differentiate between fragments with different sequences that have the same length. This may potentially lead to an underestimation of genetic diversity both within and between populations. They also need to assume that each locus has two alleles, with only one of the alleles per locus able to be amplified by PCR (Lynch and Milligan, 1994). While the universality of this assumption is questionable, treating Hj as an estimation of heterozygosity has been used recently to make population inferences from dominant data (Alexander et al, 2011). Lastly, they make the assumption that the population is in Hardy-Weinberg equilibrium. This assumption is probably valid for a locally abundant, sexually reproducing diploid insect species such as *C. humeralis*. Lynch and Milligan (1994) use the following equation to estimate heterozygosity.

$$\widehat{H}j(i) = 2\hat{q}j(i)[1 - \hat{q}j(i)] + 2Var[\hat{q}j(i)] \qquad (1)$$

This equation can be viewed as the probability that a random pair of alleles will contain one dominant allele and one null allele. Where j is the focal population and i is the locus being measured, and q is the null allele.

Estimates of population structure

The most common measure of population differentiation due to population structure is the Fixation Index or Wrights F_{ST} (1965). Wright's original definition of F_{ST} was based on the inbreeding coefficient or the probability of alleles that are identical by descent being combined in a zygote in the contemporary population. More practical

measures of F_{ST} have focused on variations in allele frequencies, such as the among population allele frequency variance (Neigel, 2002), where p is the subpopulation, and $\bar{p}$ is the average frequency of an allele in the whole population:

$$F_{ST} = \frac{Var(p)}{\bar{p}(1-\bar{p})} \quad (2)$$

However, it has been recently suggested that simple F_{ST} may be prone to underestimating the level of differentiation due to being depressed by high levels of within population variation (Meirmans, 2006). This is especially true for highly polymorphic markers such as AFLPs, because within population variance is often nearly as high as the total variance, resulting in very low between population variance that is probably not an accurate representation of the amount of gene-flow occurring between populations. Φ statistics are an analogue of F statistics that calculate the correlation of haplotypic diversity at different hierarchical scales (Excoffier, 2001), and are a flexible alternative appropriate for dominant data used by the statistical package GenAlEx 6.5 (Peakall and Smouse, 2012). Hedrik (2005) proposed the use of a standardised statistic that divides the observed Φ_{ST} by the maximum value possible given the within population variance in order to deal with the issue of high within population variation:

$$\Phi'_{ST} = \frac{\Phi_{ST}}{\Phi_{ST(MAX)}} \quad (4)$$

Both regional and population level measures of Φ_{ST} and Φ'_{ST} were calculated with GenAlEx 6.5 (Peakall and Smouse, 2012). Isolation by distance was assessed by comparing pair-wise Φ_{ST} and pair-wise geographic distance in the GenAlEx Isolation by Distance mantel test at two levels, first I included all sites and this comprised

sampling sites at very large distances from each other. Secondly, I ran a Mantel test including those seven sites that I sampled in the Arthur's Pass area.

Assignment Testing

The program STRUCTURE (Pritchard et al., 2000) is widely used for assignment tests and cluster analysis in population genetic studies. However, because of the ambiguity of the underlying genotype in dominant markers such as AFLPs, there has traditionally been some hesitation in using assignment algorithms on such genotypically depauperate datasets. In STRUCTURE 2.3.4, Falush et al. (2007) describe a new version of the Markov chain Monte Carlo algorithm that has been designed to take into account the inherent underlying ambiguity that are a part of dominant marker analysis. Falush et al. (2007), found that in a study of the whitefish *Coregonus clupeaformis* with a known ancestry, AFLP assignment tests using STRUCTURE 2.3.4 performed better than microsatellite data at assigning individuals to their correct population. While cluster analysis and searching for hybridisation were not a part of this study, Falush et al. (2007) suggest that it may be possible with STRUCTURE 2.3.4.

To carry out genotype assignment tests on *C. humeralis* I used the settings as outlined in Falush et al. (2007); I selected the USEPOPINFO option, a model that includes the population of origin as prior information. I set GENSBACK = 3, so that each individual could have detect a parent, grandparent, or great-grandparent from an alternative population, rather than the usual GENSBACK = 2. Lastly I set MIGRPRIOR = 0.001, meaning that the prior probability that an individual has pure ancestry from its designated population is 0.99.

Circuit Modeling

In addition to the model for Isolation by Distance, I created 23 resistance surfaces based on the factorial combination of the four factors of land-cover, elevation, roads, and the combination of rivers and streams (see '*Model creation and rationale*' below for more details). I then used these surfaces to calculate Isolation by Resistance, which allowed me to assess the impact of these factors on gene flow. The output model rasters were created in ARCMAP 9.3 (ESRI), which were derived from the data rich shape files found at the Landcare Research Ltd LRIS Portal for Geographic Information System (GIS) data (http://lris.scinfo.org.nz/maps/lcr/) and the Land Information New Zealand Data Service (www.data.linz.govt.nz). The raster calculator function was used to add all potential factorial levels together, apart from as discussed below. I then used CIRCUITSCAPE v3.5 (McRae et al, 2006) to produce 'pair-wise' resistance values between each of the 7 stream sites in the Arthur's Pass area that I had genetic data for. CIRCUITSCAPE also has the ability to output current maps that can provide a visual representation of the dispersal pathways present in the landscape. In this way 23 resistance matrices were generated to compare with the genetic structure of the population in the area in order to assess the impact of these factors on gene flow.

Cost Selection

The selection of the appropriate resistance values is a vital part of the use of this type of modeling framework. The resistance costs used in the models can influence the inferences that are gained from the exercise (Rayfield et al., 2010). It has been argued that the traditional methods of selecting these values, from the advice of

experts in the field, is open to subjectivity and may change from person to person (Spear et al., 2010). Instead Richardson (2012) advocates the selection of costs based on the assessment of a range of resistance values, and then using partial mantel tests to statistically assess the best fit to the genetic data. I selected four resistance cost levels of 10, 50, 100 and 500 for each of the four factors being modeled (land-cover, elevation, roads, and the combination of rivers and streams) and used partial mantel tests, calculated in the ECODIST package in R (Goslee and Urban, 2007), to estimate the correlation between the genetic distance and the resistance distance (generated by Circuitscape), controlling for the null model. The null model is an analogue of geographic distance that accounts for the fact that Circuitscape arbitrarily increases the pair-wise resistance values as the edge of the graph is approached (Amos et al., 2012). It was produced by creating a resistance surface containing a value of 1 for each pixel. A fully factorial series of hypothesis tests were conducted to be sure that factors were not interacting with each other. Calibrating the correct level of resistance costs for these potential interactions is impossible without going back and continuously manipulating the cost levels after obtaining results. This type of post testing manipulation is ill advised as it can lead to a major confirmation bias in the testing (Jonas et al., 2001). To avoid this, the highest level for each of the factors was used, with the rationale that in this way I would at least maximize the possibility that any potential interactions would be detected. Rather than finding the definitive resistance levels for each of the models factors, the goal of this exercise was to enable me to select cost values that were a good fit for the genetic data for this species in this region.

Model creation and rationale

The factor land-cover type was assigned two levels, one with high resistance and the other a null model, where the resistance was equivalent to the null model (i.e. = 1). The land-cover factor was divided into areas that are covered in native forest and those areas that are predominantly open terrain (Figure 2.1). The hypothesis that open area may act as a barrier to the active dispersal of mayflies is based on a recent study by Anderson et al. (2011) that showed an increase in genetic structure within a North American mayfly population found in areas that had a higher percentage of open areas compared to areas with a greater percentage of vegetation cover.

For elevation I wanted to test if the highest peaks were acting as genetic barriers. To test this, two levels were created, one that assigned areas higher than 1200m with a high resistance and those below with a resistance equal to the null model. 1200m was chosen because it was above the bush-line across the whole study area. The other was a null model, with no influence of ridgelines as barriers (See Table 2.3). Therefore this factor was designed to investigate whether mountain ridges act as barriers to mayfly dispersal in this area. Mayflies have generally been described as poor flyers (Malmqvist, 2000) and the brevity of their adult stage may mean that the highest ridgelines act as barriers to dispersal between catchments.

Roads were assigned three levels, a high resistance, a medium resistance, and a null model (See table 2.3). For each of these levels, highways were weighted twice the cost of gravel roads. This was done for two reasons: first, highways tend to carry more traffic and second, areas adjacent to highways tend to have increased anthropogenic impacts (Coffin, 2007). However, the greatest impact of roads on mayfly dispersal may be the barriers created by small bridges and culverts on the upstream dispersal

of adult mayflies (Blakely et al., 2006). For those resistance surfaces that combined the factors roads and rivers and streams, an additional set of maps were produced that excluded the effect of highway bridges on the fit with the genetic data. While culverts and small bridges are thought to act as barriers to along stream movement by flying stream insects (Blakely et al. 2006), this may or may not be the case for larger bridges that have a larger volume of space for individuals to move through.

Because rivers and streams are thought to act as dispersal pathways for mayflies and other stream insects (Boulton and Lake, 2008), it was important to incorporate them into the circuit models. To do this I needed to assign negative numbers to stream corridors on the raster maps in order to lower the resistance created by the other factors. This is problematic in a few ways: Firstly, Circuitscape v3.1 is unable to process negative numbers, with 1 being treated as complete conductance and anything negative treated as complete resistance by the program. Because of the model's factorial nature it was necessary that I kept the inputted values the same across all models. The consequence of this was than in a few resistance surfaces, a few stream areas became negative and thus full resistance rather than the intended full conductivity. To account for this, all negative numbers were re-assigned to 1 or full conductance, while open stream resistance was lowered compared to open land resistance. In practice, this reassignment happened in two cases; when streams run through native forest, where resistance is low, and where bridges cross forest streams. In both of these cases, the reassignment does not affect the biological question being asked, as forest is already assigned as complete conductance, and I still want to evaluate whether any resistance is created by forest roads and bridges. In this way I was able to test whether open streams added any connectivity across the open landscapes that they bisect.

Secondly, because of the high level of connectivity of streams in this landscape, if full stream conductivity was introduced into additive maps, then this factor would overwhelm the contribution of all the other factors. To deal with this, the fit of full stream conductivity to the genetic data was tested in one model, while in all other models a medium resistance compared to altitude and open areas was used. In biological terms, raster maps that have been assigned high open area resistance and a medium level of stream resistance will have streams that run through open areas having a resistance level that is far lower than the surrounding open area. This will effectively model whether or not individuals will preferentially use the stream network to disperse on if the model fit to the genetic data indicates that it is only on rare occasions that they do move over open land. Thirdly, those models that combined a null model for open areas and medium streams resistances do not address a biologically meaningful question; this is because streams are where mayflies spend the majority of their lives and evaluating them as barriers is meaningless, thus these models were not run.

Table 2.3: Factorial Design used for Resistance Modelling. (HM=High Montane Resistance, NM=Null Montane Resistance, HO=High Open Area Resistance, NO=Null Open Area Resistance, HR=High Road Resistance, MR=Medium Road Resistance, NR=Null Road Resistance, HS= High Stream Connectivity, MS=Medium Stream Connectivity, NS=Null Stream Connectivity, HB=High Bridge Resistance, NB= Null Bridge Resistance

Factor	Score	Code	Description
Mountain tops as barriers:	-	-	-
High Resistance	500	*HM*	Areas over 1200m high resistance
Null	1	*NM*	All pixels scored as 1
Land cover:			
Open Areas High Resistance	500	*HO*	All areas excluding native forest high resistance
Null	1	*NO*	All pixels scored as 1
Roads:			
High Road Resistance	300\150	*HR*	Highways scored as 300, gravel roads as 150
Medium	200\100	*MR*	Highways scored as 200, gravel roads as 100
Null	1	*NR*	All pixels scored as 1
Rivers and Streams:			
High Connectivity	-500	*HS*	Full stream connectivity
Medium Connectivity	-300	*MS*	Streams connected with medium resistance
Null	1	*NS*	All pixels scored as 1
Highway Bridge			
High Resistance	As Road	*HB*	Bridges with high resistance
No Resistance	Excluded	*NB*	Effect of Highway bridges removed

Table 2.4: Resistance Models. (HM=High Montane Resistance, NM=Null Montane Resistance, HO=High Open Area Resistance, NO=Null Open Area Resistance, HR=High Road Resistance, MR=Medium Road Resistance, NR=Null Road Resistance, HS= High Stream Connectivity, MS=Medium Stream Connectivity, NS=Null Stream Connectivity, HB=High Bridge Resistance, NB= Null Bridge Resistance

Model Name	Factorial Levels
Res Model 1	*NM/NO/NR/NS*
Res Model 2	*HM/HO/HR/MS/HB*
Res Model 3	*HM/NO/NR/NS*
Res model 4	*NM/HO/NR/NS*
Res Model 5	*NM/NO/HR/NS*
Res Model 6	*HM/HO/HR/HS*
Res Model 7	*HM/HO/MR/MS/HB*
Res Model 8	*NM/NO/MR/NS*
Res Model 9	*NM/HO/HR/MS/HB*
Res Model 10	*HM/HO/NR/MS*
Res Model 11	*HM/HO/HR/NS*
Res Model 12	*HM/HO/NR/NS*
Res Model 13	*HM/NO/HR/NS*
Res Model 14	*NM/HO/HR/NS*
Res Model 15	*NM/HO/NR/MS*
Res Model 16	*HM/HO/MR/NS*
Res Model 17	*NM/HO/MR/MS/HB*
Res Model 18	*HM/NO/MR/NS*
Res Model 19	*NM/HO/MR/NS*
Res Model 20	*HM/HO/HR/MS/NB*
Res Model 21	*NM/HO/HR/MS/NB*
Res Model 22	*NM/HO/MR/MS/NB*
Res Model 23	*HM/HO/MR/MS/NB*

Model Evaluation

Pearson ranked partial Mantel tests (1000 permutations) were run on each of the competing models using the '*ecodist*' package in the R statistical environment (Goslee and Urban, 2007). Partial Mantel tests were used to find the correlation between two pair-wise distance matrices while controlling for the effects of a third as proposed by Smouse et al. (1986). For population genetic studies utilizing resistance matrices generated in Circuitscape, we are looking to find the amount of correlation between the pair-wise genetic matrix and the landscape features codified within the resistance matrix while controlling for a null resistance matrix that is a corrected analogue of geographic distance (as explained above).

There has been some debate recently concerning the efficacy of the partial Mantel test, however it appears that the major criticism is a loss of statistical power and an increased likelihood of committing a type 2 error rather than of an increased potential to make spurious inferences (Richardson, 2012). However in certain circumstances there may be a slightly increased chance of finding a relationship that is not there (type 1 error), especially when there are severe outliers in the dataset (Legendre, 2000). To deal with outliers, Legendre and Fortin (2010) used an alternative permutation method, this method involves permuting the residuals of the null regression model to account for this, and otherwise they found that permuting the raw scores provided low type 1 error and good power. I tested whether there were any such extreme outliers using the outlier.ranking command in the Data Mining package (DMwR) in the R statistical suite (Trogo, 2010), finding no outliers that were impacting on the dataset (See Results). Lastly, Cushman et al., (2013) suggest that there is a much reduced chance of committing a type 1 error if, when evaluating the fit of resistance models to genetic data, comparisons are based on the partial mantel

correlation coefficient rather than the p-value which is more prone to influence by correlated alternative hypotheses.

I carried out 23 Partial Mantel tests based on my previously outlined factorial design, using Smouse et al's 1987 permutation of the raw scores of matrix A. Rather than using the Bonferroni method for correcting for multiple tests of significance, I used the less conservative Benjamani and Hochberg (1995) B-H correction. It is important to point out that partial Mantel tests do not produce coefficients of determination as a general linear model would, but rather they produce a correlation coefficient and therefore can only be used to make statements regarding the degree to which the model fits the data and not regarding the level to which landscape variables determines the pattern of genetic data (Legendre and Fortin, 2010). The partial correlation coefficients that were produced from this series of partial Mantel tests were used to examine two things: which individual factor contributed most to the model with the best fit with the genetic data, and which combination of factors gave the best model fit. By ranking the models that best fit the genetic data I was able to evaluate which of the landscape features chosen had the most effect on gene-flow in the Arthur's Pass area.

Chapter 3
Results

I successfully generated AFLP genotypes for 162 individual *C. humeralis* nymphs from my 10 streams. These genotypes consisted of 209 loci generated from two selective primers. These loci were highly polymorphic, with 207 of the 209 fragments (99%) being useful for population genetic analysis. From the individuals that were randomly selected to be replicated there was a final mismatch error rate of 6.05% for the selective primers pairs used.

Genetic Diversity

Across all individuals genotyped, there was an average of 59.7 fragments scored per individual. There were on average 56.5% polymorphic loci per population (PLP). However there was a large variety between the 10 sites that were sampled. Two sites (Pylon Gully: 35.4%; Kaituna: 30.1%) had very low PLP compared to the other sites sampled (An average of 62.5% PLP excluding Pylon Gully and Kaituna) (Table 3.1). Expected heterozygosity (Hj) followed a similar pattern, with the Pylon Gully and Kaituna sites having considerably lower Hj than the other populations (Table 3.1).

Table 3.1: Genetic diversity for *Coloboriscus humeralis* larvae collected from 10 stream sites across New Zealand. N is the number of analyzed samples; PLP is the polymorphic loci per population (at the 5% level); Hj is the expected heterozygosity.

Site/Population	N	PLP (%)	Hj (±SE)
Castle Hill	16	42.1	0.160 (0.013)
Cheeseman	9	65.1	0.171 (0.009)
Craigieburn	13	75.6	0.220 (0.011)
Pylon Gully	16	35.4	0.126 (0.011)
Peacock	23	51.7	0.186 (0.012)
Moss	21	61.7	0.208 (0.013)
Waimak Bend	22	64.1	0.204 (0.012)
Kaituna	12	30.1	0.116 (0.012)
Te Oka Bay	9	69.9	0.249 (0.013)
Northland	19	69.4	0.242 (0.013)

Population Structure

At all geographic scales, there was a high level of genetic structure between the streams. Between all populations including the out-group in Northland, there was relatively high genetic differentiation (ΦPT = 0.209, p = 0.001; Φ'PT (corrected for within population variation) =0.229, p=0.001). Between the three geographic regions of Northland, Banks Peninsula, and the Arthur's Pass area located in the South Island's Southern Alps, there was a significant amount of genetic differentiation (ΦRT = 0.024 p = 0.015).

All pairs of individual populations had pair-wise ΦPT values with p-values below the standard α level of 0.05 (See Table 3.2). However, when the influence of multiple tests is taken into account using the B-Y correction as proposed by Narum (2006) for conservation genetics analyses, the α level becomes 0.0137. This means that two population pairs, Te Oka Bay and Craigieburn, and Te Oka Bay and Waimak Bend can no longer be classed as significantly genetically distinct. However, the low number of usable genotypes I was able to obtain from Te Oka Bay, coupled with the large distance between these streams in the context of a population that shows high levels of genetic structure at very small geographic scales, suggests that these two results are more likely to be artefacts of low sample size rather than evidence of long distance dispersal between the Southern Alps and Banks Peninsula. All other pair-wise comparisons of genetic differentiation remained significant even at this lower α level.

Pair-wise ΦPT ranged from 0.065 (p=0.011) between Moss and Peacock, two streams divided by only 1.7 kilometres, to 0.331 between Craigieburn and Moss, two streams separated by 16.7 kilometres (See Table 2.1.). As expected, the out-group, Northland

(An average of 789.25 kilometres from the other sites, and located on the North Island), had high levels of pair-wise genetic differentiation with the majority of the other sample sites (Table 3.2.).

Assignment testing in Structure (Pritchard et al., 2000, Falush et al., 2007) was able to assign a substantial proportion of the genotypes to either the putative population of origin or a geographically close neighbouring stream for all but one of the populations (See Table 3.3.). Kaituna Valley stream, located on Banks Peninsula was given an assignment of 0.903 to Pylon Gully, a stream located 121.1 kilometres away in the Cass area.

TABLE 3.2: Pair-wise ΦPT: The bottom left quadrant details estimated molecular variance between individual populations (ΦPT), while the top right gives the probability values (p). * denotes significance at the B-Y corrected level of 0.0137

Sites	**Castle Hill**	**Cheeseman**	**Craigieburn**	**Pylon Gully**	**Peacock**	**Moss**	**Waimak bend**	**Kaituna**	**Te Oka**	**Northland**
Castle Hill	---------	0.001*	0.001*	0.001*	0.001*	0.001*	0.001*	0.001*	0.001*	0.001*
Cheeseman	0.163	---------	0.013*	0.001*	0.002*	0.001*	0.001*	0.001*	0.001*	0.001*
Craigieburn	0.185	0.101	---------	0.001*	0.001*	0.001*	0.005*	0.002*	0.019	0.001*
Pylon Gully	0.230	0.224	0.174	---------	0.008*	0.001*	0.001*	0.003*	0.001*	0.001*
Peacock	0.206	0.249	0.174	0.096	---------	0.011*	0.002*	0.001*	0.006*	0.001*
Moss	0.294	0.331	0.201	0.226	0.065	---------	0.012*	0.001*	0.005*	0.001*
Waimak bend	0.257	0.242	0.120	0.212	0.111	0.075	---------	0.001*	0.033	0.001*
Kaituna	0.225	0.184	0.170	0.124	0.180	0.274	0.228	---------	0.001*	0.001*
Te Oka	0.273	0.251	0.098	0.237	0.118	0.108	0.078	0.216	--------	0.001*
Northland	0.225	0.182	0.122	0.250	0.191	0.215	0.148	0.247	0.139	--------

Table 3.3: Assignment test. Proportion of membership from predefined population in each of the 10 assigned sites using STRUCTURE 2.3.4 (Pritchard et al. 2000). Given assigned sites: The sites on the furthest left are where the individuals were collected; along the top is the proportion of the genotyped individuals that were assigned to each site.

Sites	Castle Hill	Cheeseman	Craigieburn	Pylon Gully	Peacock	Moss	Waimak Bend	Kaituna	Te Oka	Northland
Castle Hill	0.858	0.125	0	0	0	0	0	0.016	0	0
Cheeseman	0	0.900	0	0.006	0.002	0	0.061	0.027	0.003	0
Craigieburn	0	0.308	0.267	0	0	0	0.416	0	0.008	0
Pylon Gully	0	0	0	0.613	0.387	0	0	0	0	0
Peacock	0	0	0.082	0.326	0.469	0.016	0	0.034	0.073	0
Moss	0	0	0.003	0.001	0.191	0.697	0.102	0.005	0.002	0
Waimak Bend	0	0	0.416	0	0.004	0	0.533	0.046	0	0
Kaituna	0	0	0	0.903	0.002	0	0	0.094	0	0
Te Oka	0	0	0.014	0	0.001	0.001	0.267	0.010	0.706	0
Northland	0	0.011	0.002	0.002	0	0.014	0.007	0.022	0.002	0.940

Table 3.4: A list of Individuals with inferred ancestry from other than the stream they were collected from. As calculated by the assignment test function in STRUCTURE 2.3.4 (Pritchard et. al., 2000).

Probability of being from population the same as collected from.			**Probability of being from other population.**	
Collected from:	**Label**	**Probability**	**Inferred ancestral site:**	**Probability**
Castle Hill	Ch1	0	**Cheeseman**	1
Castle Hill	Ch2	0	**Cheeseman**	1
Castle Hill	Ch7	0	**Te Oka Bay**	1
Cheeseman	Cm18	0.206	**Waimak Bend**	0.47
Cheeseman	Cm19	0.003	**Waimak Bend**	0.438
Craigieburn	CR1	0	**Cheeseman**	1
Craigieburn	Cr10	0	**Waimak Bend**	1
Craigieburn	Cr11	0	**Waimak Bend**	1
Craigieburn	Cr12	0	**Waimak Bend**	1
Craigieburn	Cr17	0	**Waimak Bend**	1
Craigieburn	Cr19	0	**Waimak Bend**	1
Craigieburn	CR3	0	**Cheeseman**	1
Craigieburn	CR5	0	**Cheeseman**	1
Craigieburn	CR7	0	**Cheeseman**	1
Craigieburn	Cr9	0.011	**Waimak Bend**	1
Pylon Gully	Py10	0	**Peacock**	1
Pylon Gully	Py12	0	**Peacock**	1
Pylon Gully	Py18	0.251	**Peacock**	0.495
Pylon Gully	Py2	0	**Peacock**	0.996
Pylon Gully	Py20	0	**Peacock**	0.987
Pylon Gully	Py21	0	**Peacock**	1
Pylon Gully	Py6	0	**Peacock**	1
Peacock	Pa1	0	**Pylon Gully**	1
Peacock	Pa10	0.049	**Moss**	0.44
Peacock	Pa11	0	**Pylon Gully**	1
Peacock	Pa12	0	**Pylon Gully**	1
Peacock	Pa16	0.049	**Pylon Gully**	0.66
Peacock	Pa18	0	**Pylon Gully**	1
Peacock	Pa19	0	**Pylon Gully**	1
Peacock	Pa20	0	**Pylon Gully**	0.998
Peacock	Pa24	0	**Craigieburn**	0.51
Peacock	Pa4	0	**Te Oka Bay**	0.747
Peacock	Pa5	0	**Pylon Gully**	1
Peacock	Pa6	0	**Craigieburn**	0.529
Peacock	Pa9	0	**Craigieburn**	0.532

Table 3.4 (Cont.): A list of Individuals with inferred ancestry from other than the stream they were collected from. As calculated by the assignment test function in STRUCTURE 2.34 (Pritchard et. al., 2000).

Probability of being from population the same as collected from.			**Probability of being from other population.**	
Collected from:	**Label**	**Probability**	**Inferred ancestral site:**	**Probability**
Moss	M12	0	**Waimak Bend**	1
Moss	M14	0	**Peacock**	0.989
Moss	M15	0	**Peacock**	1
Moss	M16	0	**Peacock**	1
Moss	M18	0	**Peacock**	1
Moss	M6	0.367	**Waimak Bend**	0
Moss	M8	0	**Waimak Bend**	0.985
Waimak Bend	Wb10	0	**Craigieburn**	0.766
Waimak Bend	Wb14	0	**Craigieburn**	0.999
Waimak Bend	Wb18	0	**Craigieburn**	0.762
Waimak Bend	Wb13	0	**Craigieburn**	1
Waimak Bend	Wb16	0	**Craigieburn**	1
Waimak Bend	Wb17	0	**Craigieburn**	0.849
Waimak Bend	Wb21	0	**Kaituna**	0.989
Waimak Bend	Wb22	0	**Craigieburn**	1
Waimak Bend	Wb23	0	**Craigieburn**	1
Waimak Bend	Wb24	0	**Craigieburn**	1
Waimak Bend	Wb7	0.272	**Craigieburn**	0.245
Waimak Bend	Wb9	0	**Craigieburn**	1
Kaituna	Kt10	0	**Pylon Gully**	1
Kaituna	Kt11	0	**Pylon Gully**	1
Kaituna	Kt12	0	**Pylon Gully**	1
Kaituna	Kt13	0	**Pylon Gully**	1
Kaituna	Kt14	0	**Pylon Gully**	0.837
Kaituna	Kt15	0	**Pylon Gully**	1
Kaituna	Kt2	0.117	**Pylon Gully**	0.538
Kaituna	Kt3	0	**Pylon Gully**	1
Kaituna	Kt4	0	**Pylon Gully**	1
Kaituna	Kt6	0	**Pylon Gully**	1
Kaituna	Kt7	0	**Pylon Gully**	1
Kaituna	Kt8	0	**Pylon Gully**	0.953
Te Oka	To20	0	**Pylon Gully**	1
Te Oka	To21	0	**Pylon Gully**	1
Te Oka	To22	0.101	**Waimak Bend**	0.637
Northland	Nl17	0.03	**Moss**	0.825

There were a total of 52 individuals that STRUCTURE (Pritchard, et al., 2000) assigned to a stream site other than the one from which it was collected with a probability over 95% (Table 3.4). Of these 52 individuals, 10 were assigned to Pylon Gully that were collected on Banks Peninsula at Kaituna Valley. Out of the remaining 42 putative immigrants from the last three generations, 25 or 59.52% originated from an adjacent population, no more than 9 kilometres away. A further 13 individuals may indicate recent gene-flow between Craigieburn and Waimak Bend streams, two sites that are located 15.63 kilometres apart. There was also several inferred migrants between Peacock stream and Pylon Gully, with the two streams being separated by 8.3 kilometres, and predominately open agricultural area.

Isolation by Distance

At the largest regional scale, there was no discernible effect of isolation by distance ($r^2 = 0.243$, $p = 0.970$, Figure 3.1). However, the very large geographic distance between Northland and the Cass area, coupled with the significant pattern of Isolation by Distance (IBD) when only considering the seven sites in the Cass/Arthur's Pass area ($r^2 = 0.753$, $p = 0.001$, Figure 3.2), may indicate that at the largest regional scale, homoplasy is responsible for a lessening of the actual variation between these regions. Furthermore, as the r^2 value for IBD is significant when looking at the Cass/Arthur's Pass area, having the most individuals and therefore the most power to statistically infer a spatial pattern, is probably more reliable in describing the movement of genes in this species.

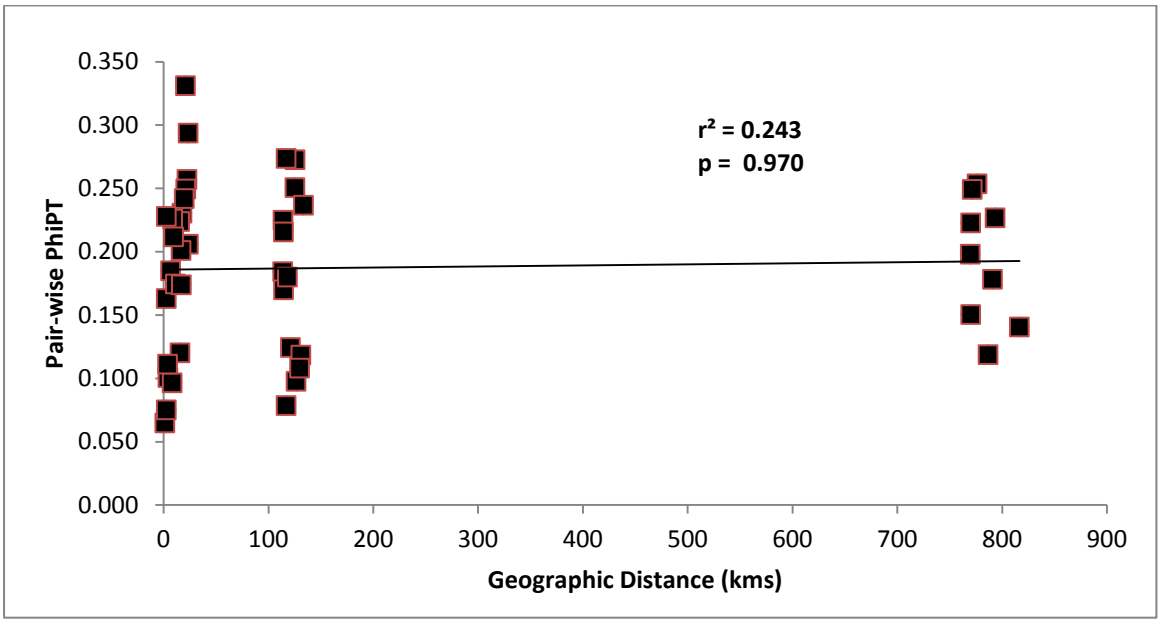

Figure 3.1: Isolation by Distance for all study sites. Each point represents the pair-wise genetic distance (ΦPT) between each population of the New Zealand mayfly, *Coloburiscus humeralis,* and the relationship to geographic distance. For this analysis all study sites were included. Significance was assessed using a Mantel test using Arlequin 3.0 (Excoffier, et al, 2005).

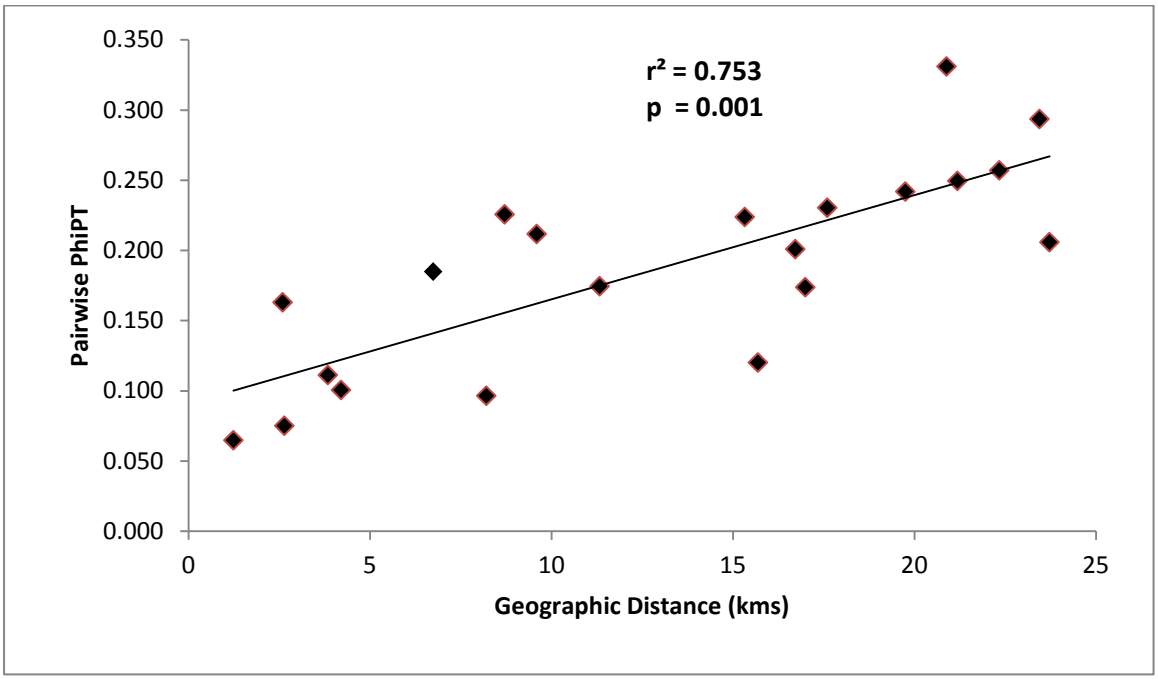

Figure 3.2: Isolation by Distance for the Cass/Arthur's Pass area. Each point represents the pair-wise genetic distance (ΦPT) between each population of the New Zealand mayfly, *Coloburiscus humeralis,* and the relationship to geographic distance. For this analysis only the 7 sites in the Cass Arthur's Pass area were included. Significance was assessed using a Mantel test using Arlequin 3.0 (Excoffier, et al, 2005).

Circuit Modelling

Cost Parameterisation

Cost parameterisation showed that the highest resistance cost level tested generated the best fit for the genetic data (See Table 3.5). This was especially so when open area resistance was given a cost of 500, this provided a partial correlation coefficient (PCC) of 0.534 ($p = 0.002$). While no higher cost levels were tested, there is a considerable pattern of diminishing returns when comparing the difference in partial correlation between the open area cost levels of 50-100 (Δ F2-F3 PCC = 0.111) compared with the difference between the substantially larger cost difference between 100 and 500 (Δ F3-F4 PCC = 0.005). This would appear to be good evidence that any higher cost levels would be asymptotic. While the other factors did not have as great a fit to the genetic data as open areas, the top cost level was used for the reasons outlined in the methods. Because of this the final full factorial resistance modelling design was run using the highest level for all factors (Table 3.5).

Table 3.5: Resistance Cost testing Results. Cost parameterisation for the 4 main factors, using partial Mantel tests to remove the influence of distance using the corrected null model analogue of Euclidean distance as calculated in the ecodist package in R (Goslee and Urban, 2007).

Factor	Level Name	Level Cost	Partial Correlation	Raw p-value
Open Area	F1	10	0.393	0.047
	F2	50	0.418	0.048
	F3	100	0.529	0.007
	F4	500	0.534	0.002
Mountains	M1	10	-0.033	0.574
	M2	50	-0.037	0.568
	M3	100	-0.047	0.595
	M4	500	-0.049	0.58
Roads	R1	5\10	-0.292	0.864
	R2	15\30	-0.273	0.839
	R3	30\60	-0.224	0.774
	R4	150\300	-0.223	0.791
Stream	S1	10	0.001	0.511
Connectance	S2	50	0.115	0.322
	S3	100	0.131	0.291
	S4	500	0.0195	0.505

Outliers

There were no pair-wise genetic distance (ΦPT) values that were considered as being significant outliers, that might increase the likelihood of committing a type 1 error (See Table 3.6). Because of this, partial Mantel test were run on the resistance models using the raw ΦPT values, and not the residuals of a regression equation between ΦPT and the null model (Legendre, 2011).

Circuit Modelling Results

Out of the 23 landscape-resistance hypotheses that were tested, 11 were significantly supported by the genetic data when controlling for distance ($p \leq 0.05$). Another 5 models were just over the significance threshold when a B-H correction for multiple testing had been performed (corrected $p = 0.051$). These well fitting models all had one thing in common: all had a high open area resistance. On the other hand, all models that incorporated the no open area resistance level had very bad model fits with the genetic data. Indeed, all of these hypotheses had either negative or very low partial correlation coefficients (see Table 3.7).

Of those hypotheses that estimated the model fit of one factor at a high level, while maintaining the rest at null resistance, only Model 4 with high open area resistance provided a significant fit to the genetic data ($p = 0.037$, Figure 3.3). By including the effects of mountain tops as barriers ($r^2 = -0.047$, $p = 0.692$), roads ($r^2 = -0.223$, $p = 0.797$), and full stream connectance ($r^2 = 0.0195$, $p = 0.591$), while all others were held at the null level, there was little to no fit to the genetic data. While some of the additive models remained significant when these factors were included, the fact that model 4 was the top ranked model, and solely represents a high level of open area resistance with no resistance for dispersers travelling across forested areas, would suggest that no interaction between factors significantly increases the fit of the

model. When a partial Mantel test was run that measured the relationship between the pair-wise ΦPT and Euclidean distance with the effect of model 4 partialled out there was no fit to the genetic data ($r^2 = 0.0262$, $p = 0.128$). This suggests that there is an effect of open area resistance even when the strong effect of isolation by distance is accounted for. These results suggest that areas of land that are not covered in native forest and are predominately open provide a significant barrier to the gene-flow of *Coloburiscus humeralis* in the Arthur's pass region.

TABLE 3.6: Outlier Detection. The bottom left quadrant details estimated molecular variance between individual populations (ΦPT), while the top right gives the probability that a pair-wise value is an outlier, in brackets is the outlier ranking of each ΦPT value (Trogo, 2007). Any value over 0.950 has a significant probability of being an outlier.

Sites	**Castle Hill**	**Cheeseman**	**Craigieburn**	**Pylon Gully**	**Peacock**	**Moss**	**Waimak bend**
Castle Hill	---------	0.466 (6)	0.500 (3)	0.333 (13)	0.000 (18)	0.636 (1)	0.333 (14)
Cheeseman	0.163	---------	0.429 (7)	0.000 (19)	0.333 (15)	0.636 (2)	0.333 (16)
Craigieburn	0.185	0.101	---------	0.466 (4)	0.466 (5)	0.000 (20)	0.429 (8)
Pylon Gully	0.230	0.224	0.174	---------	0.429 (9)	0.000 (21)	0.333 (17)
Peacock	0.206	0.249	0.174	0.096	---------	0.429 (10)	0.429 (11)
Moss	0.294	0.331	0.201	0.226	0.065	---------	0.429 (12)
Waimak bend	0.257	0.242	0.120	0.212	0.111	0.075	---------

Table 3.7: Resistance Models. Results from partial Mantel tests, where the effect of distance is partialled out, leaving the relationship between the resistance model as detailed below and the genetic data. Factorial level codes outlined in Methods. * denotes significance.

Model Name	Factorial Levels	Rank	Partial correlation	p-values (BH corrected)
Res model 4	*NM/HO/NR/NS*	1	0.534	0.038*
Res Model 12	*HM/HO/NR/NS*	2	0.521	0.038*
Res Model 20	*NM/HO/MR/NS*	3	0.523	0.038*
Res Model 14	*NM/HO/HR/NS*	4	0.513	0.039*
Res Model 15	*NM/HO/NR/MS*	5	0.505	0.040*
Res Model 18	*NM/HO/MR/MS/HB*	6	0.490	0.040*
Res Model 9	*NM/HO/HR/MS/HB*	7	0.488	0.039*
Res Model 23	*NM/HO/MR/MS/NB*	8	0.487	0.038*
Res Model 22	*NM/HO/HR/MS/NB*	9	0.487	0.038*
Res model 17	*HM/HO/MR/NS*	10	0.461	0.040*
Res Model 11	*HM/HO/HR/NS*	11	0.460	0.040*
Res Model 10	*HM/HO/NR/MS*	12	0.418	0.051
Res Model 7	*HM/HO/MR/MS/HB*	13	0.407	0.051
Res Model 28	*HM/HO/MR/MS/NB*	14	0.406	0.051
Res Model 2	*HM/HO/HR/MS/HB*	15	0.406	0.051
Res Model 21	*HM/HO/HR/MS/NB*	16	0.404	0.051
Res Model 1	*NM/NO/NR/NS*	17	0.269	0.140
Res Model 6	*HM/HO/HR/HS*	18	0.020	0.591
Res Model 3	*HM/NO/NR/NS*	19	-0.047	0.692
Res Model 19	*HM/NO/MR/NS*	20	-0.183	0.797
Res Model 13	*HM/NO/HR/NS*	21	-0.203	0.797
Res model 8	*NM/NO/MR/NS*	22	-0.225	0.797
Res Model 5	*NM/NO/HR/NS*	23	-0.223	0.797

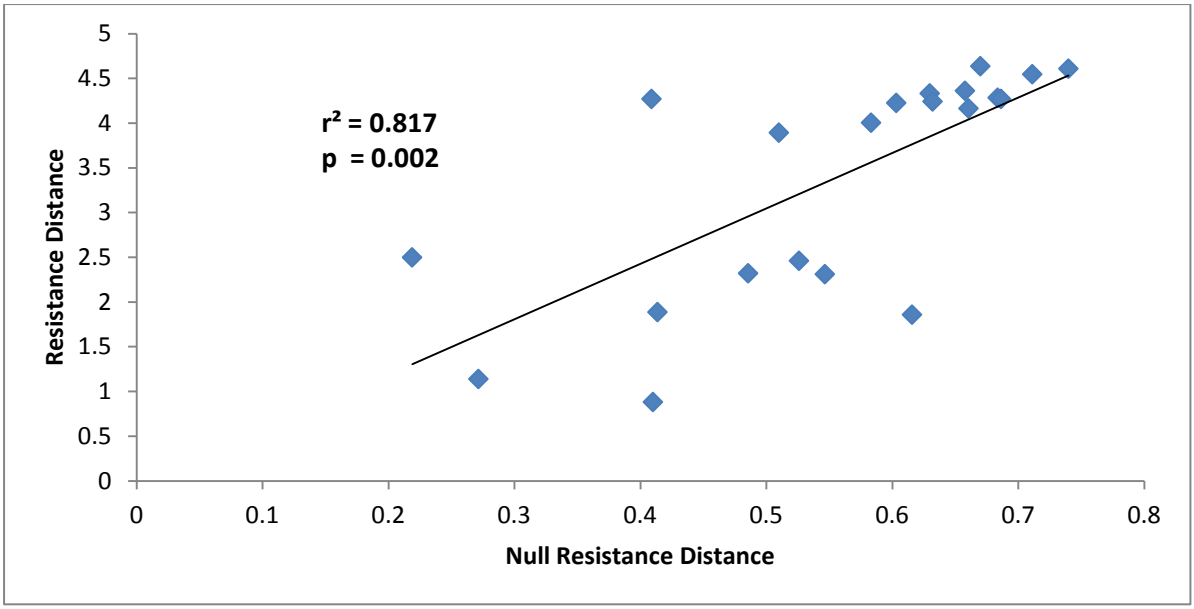

Figure 3.3: Isolation by Resistance for Model 4. Each point represents the relationship between resistance distance (generated from model 4 using the programme Circuitscape (McRae, 2006)) and the null resistance distance (an analogue of geographical distance that has been corrected for the influence of edge effects, as generated by Circuitscape (McRae, 2006)).

Chapter 4

Discussion

Outline

In this study of ten forested stream populations of the mayfly *Coloburiscus humeralis* I wanted to use AFLP techniques to determine the genetic relationships between these populations. I predicted that there would be higher genetic differentiation between populations than was measured in these previous using allozymes (Hogg et al, 2003; Morris, 2005), because the use of slow mutating allozymes was likely to underestimate the population structure present in this species (Meudt and Clarke, 2006). I also hypothesised that due to the affinity of this species to forested streams (Harding and Winterbourn, 1996), that dispersal may be occurring more frequently in forested areas than across open landscapes.

Genetic diversity

Seven out of the ten populations in this study had a much higher level of gene diversity than the remaining three (both Percentage of Polymorphism and Expected Heterozygosity). High genetic diversity is often associated with robust and viable populations, while measuring genetic diversity may provide insight into the effects of complicated environmental stressors. For example, in a meta-analysis, Bickman et al,

(2000) found that those populations exposed to high levels of environmental contamination consistently had lower levels of genetic diversity than those that were found in uncontaminated ecosystems. This suggests that these seven populations may be healthier than those in populations with lower diversity. Conversely, three populations (Kaituna Valley, Castle Hill and Pylon Gully) had low measures of gene diversity. While making direct conclusions regarding inbreeding using dominant markers can be problematic, in general low genetic diversity may correlate with lowered overall fitness (Alexander et al, 2011). Recent evidence using the estuarine crustacean, *Americamysis bahia* suggests that low genetic diversity as assessed by AFLPs can have distinct and measurable effects on population fitness in an experimental setting (Markert et al, 2010). Other studies have found a significant correlation between genetic diversity found in AFLP data sets and population fitness (Hufford et al, 2012). In populations that are isolated and that may have undergone a population contraction in recent years, genetic drift may result in the fixation of detrimental alleles, resulting in lowered genetic diversity and potentially inbreeding depression (Hendriks and Kalinowski, 2000). Furthermore, a previous study by Hogg et al, (2002) using allozymes found significant levels of inbreeding in their populations of *Coloburiscus humeralis*. My results would seem to be at odds with their finding of no genetic differentiation at any scale considering the large local population sizes observed for this species. It therefore may be that the three streams in this study which had much lower genetic diversity than the others is an indication that these populations are less fit and more susceptible to local extinction due to pulsed or chronic disturbances. Interestingly, all three of these streams were separated from other *Coloburiscus* populations by open non-forested land, which supports the landscape genetic results discussed later in this chapter. Anecdotally, while in most streams the collection of 30 *Coloburiscus humeralis* larvae was

accomplished on average in around 5 kick-net samples, in Kaituna Valley and in particular Pylon Gully it took considerably longer to collect that many insects. While this does not provide conclusive evidence that these three populations were smaller than others, combined with their physical isolation and low genetic diversity it may be that these populations are less robust than others in the study.

Population Genetics

A strong pattern of genetic differentiation was clearly evident among the *Coloburiscus humeralis* populations in this study. Nearly all populations were highly genetically distinct and there was little evidence of recent inter-stream dispersal. This pattern of genetic differentiation was present even for the pair of streams that were both geographically the closest to one another in this study and connected by beech forest (*Nothofagus spp.)*. These two streams (Peacock and Moss Streams) were only separated in a direct line by 1.2 km, and while the differentiation between the two sites (Φ_{PT} =0.065, p=0.011) was low compared to some of the other pair-wise measures it was still significant (See Table 3.2). These streams both flow into the Waimakariri River and the significant genetic differentiation is a strong indicator of dispersal limitation in this species. In both the streams that were inside continuous forest and those that were outside in isolated forest fragments there was significant pair-wise genetic differentiation. This may indicate that regardless of the level of forest cover between streams that in this species inter-stream dispersal is a rare event. This finding supports previous work that suggested that the majority of adult stream insects will need to disperse upstream to counteract the population loss due

to downstream drift (Müller, 1982). However, more sampling that includes multiple sites within the same stream would have to be done to confirm this.

In the Arthur's Pass area, there was a strong and significant pattern of isolation by distance. That is, as the geographic distance between two sites increased so did the amount of genetic differentiation between the sites. Isolation by distance is a common finding in many population genetics studies, it has been found in many species across a range of taxa, including the global human population (Relethford, 2004), a variety of plants (Moyle, 2006), and a number of insect species, including a recent study on a stream dwelling blackflies (Simuliidae) (Finn et al, 2006). At small spatial scales isolation by distance can be seen as the null hypothesis that more complete models incorporating landscape heterogeneity are compared against. A strong pattern of isolation by distance may be an indication that while dispersal events may be rare between these streams, they are still providing an important role in generating the population genetic structure that has been observed (Slatkin, 1993). This is because, to create such a pattern adults must be dispersing more often to neighbouring streams than to streams that are further afield in what might be described as a 'stepping stone' model of dispersal (MacArthur and Wilson, 1967). In contrast, if there were insurmountable dispersal barriers and there was no recent gene flow between the populations in the study, we would not expect a pattern of isolation by distance (Keyghobadi et al, 2005).

As was the case in the streams sampled in Arthur's Pass National Park and its surroundings, the two populations from the two sites on Banks Peninsula were also significantly genetically distinct from one another (Φ=0.216, p=0.001). Banks Peninsula is what remains of two extinct volcanoes, creating the Lyttelton and Akaroa harbours today. The average height of the Akaroa erosion caldera that

separates the sites sampled in this study is ~760m a.s.l. (Stipp and McDowell, 1974). My two streams (Kaituna Valley and Te Oka Bay) were separated by 14.1 kilometres, and while the Te Oka stream is located in a steep-sided bay on the southern side of the peninsula, Kaituna valley is wide and overlooks Lake Ellesmere (see figure 2.2). As in many parts of New Zealand, Banks Peninsula has undergone extensive deforestation since European settlement and today only fragments of forest remain (Wood and Pawson, 2008). Suitable habitat for many New Zealand mayflies is often associated with intact native forests or native tussocklands. Harding and Winterbourn (1995) found that stream invertebrate diversity was considerably higher in waterways running through native forest than those running through plantation forests or developed pastures. It may be that the lack of suitable streams or 'habitat islands' between these two populations are preventing adequate gene flow via a stepping stone model (MacArthur and Wilson, 1967).

In recent years there have been a number of studies using rapidly mutating markers to assess the population genetic structure of a number of mayfly species. In the endangered European alpine mayfly, *Ameletus inopinatus, is* found at over 600m a.s.l. in the British Isles and Central Europe, but it occurs at lower altitudes in colder northern Eurasian streams. Taubman et al. (2011) used microsatellites to determine that there was little dispersal in this mayfly occurring across areas characterised by warmer climactic conditions, but substantial movement within these habitat islands While a study of the North American desert dwelling mayfly species, *Callibaetis americanus,* found that there was significant genetic differentiation between desert springs within the same valley, suggesting little dispersal between these springs (Stutz et al. 2010). Comparing the population structure between a lotic and a lentic specialist species of mayfly of the same genus, Drotz et al. (2012) found that there

was a considerably higher genetic differentiation between sites for the stream dwelling species than for the species that lives in the margins of lakes; suggesting that the unpredictable nature of streams compared to the more stable lake environment was contributing to this pattern. In New Zealand, Smith et al. (2006) found a high level of genetic differentiation between populations of *Acanthophlebia cruentata* (a widely distributed North Island mayfly species). This demonstrates that dispersal limitation may be a common trait in a number of freshwater mayflies. In contrast to these studies, the Australian mayfly *Ulmerophlebia* sp. was shown to have low levels of genetic differentiation between nearby streams, indicating a frequent exchange of individuals between streams, and it was only when sites were >15 km apart that there was any genetic differentiation (Young et al. 2012). While most of these studies appear to find strong between region genetic variation, within region measurements of variation are more varied. These variable results from different studies may reflect the diversity of dispersal abilities present in any large taxonomic group such as the Ephemeroptera, and differences in climate, landscape features, and evolutionary history in different regions. Thus, it is important to be aware that a pattern of very limited gene flow between streams may not be applicable to all mayfly species.

My result of strong genetic differentiation at small spatial scales differs from patterns found in the previous studies on this mayfly species (Hogg et al., 2002; Morris., 2005). Hogg et al. (2005) found that in populations selected from across New Zealand there was no relationship between genetic distance and geographical distance. They found that in sites separated by hundreds of kilometres and located on North and South Islands that there was no pattern of genetic differentiation, nor was there a difference between populations in sites separated by 1.5 kilometres in the

Waikato Region of the North Island. However, both Hogg et al. (2002) and Morris (2005) used allozyme techniques, which do not provide a sufficient resolution to provide information on the movement of genes in the most recent generations and therefore are less likely to resolve modern patterns of dispersal (Meudt and Clark, 2007). For example, a comparison between allozymes and microsatellites in populations of Russian Chum Salmon, found that microsatellites had a far greater ability to resolve recent gene flow (Rubtsova et al. 2008). In my study there were also two sites that had a similar distance apart as those sampled in the Hogg et al, (2002) paper; my Northland site and Te Oka bay were separated by 828 kilometres and were significantly genetically different from one another ($\Phi PT=0.139$, $p=0.001$). While he found little genetic variation within regions, Morris (2005) found significant between region genetic differentiation in *Coloburiscus humeralis*, and also found that one site on the West Coast was significantly different from the other streams in the region. Hogg et al. suggested that this may reflect a dispersal barrier caused by the presence of a large glacial mass between the two sites during the last glacial maximum. While allozyme markers are often not sufficient to resolve recent patterns of gene flow, Bossart and Prowell (1998) suggest they may be suitable for resolving historical patterns of vicariance. Combined with the fact that this species has such a wide distribution across New Zealand, this may suggest that landscape changes (e.g., forest clearing) have led to restricted dispersal in this species.

There was a significant level of genetic differentiation between the Arthur's Pass, Banks Peninsula and the Northland regions, suggesting that little dispersal was occurring between these geographically distinct regions ($\Phi RT=0.03$, $p=0.009$). However, while there was a strong pattern of isolation by distance in the Arthur's Pass region, there appears to be no relationship between distance and genetic

similarity at the regional level, as is demonstrated by the lack of any pattern of isolation by distance when all study sites are taken into account (Fig 3.1). Because measures of isolation by distance assume migration-drift equilibrium, Rousset (1997) suggest that they are only effective at small geographic scales. This suggests that the finding of isolation by distance in the Arthur's Pass area is the one that should be given the most weight.

Furthermore, several measures of pair-wise genetic differentiation found between regions were not considerably higher than between populations in the same region. Homoplasy, caused by the co-migration of two or more unique fragments at the same 'locus' during electrophoresis (e.g. when independent mutations cause the loss of the same fragment), often results in an underestimation of genetic variation when AFLPs are used as markers; this is especially true as populations become more genetically dissimilar (Meudt and Clarke, 2007). Because homoplasy masks genetic differentiation between populations that are likely to have been separated over long time periods due to the large geographic distances separating them, this means that the similar levels of genetic differentiation within and among regions in this study is to be expected. Lastly, the high levels of genetic differentiation at local scales found in my study suggests that dispersal across large inter-regional distances is unlikely or very rare, suggesting that the genetic similarity found between Te Oka Bay and the two sites in the Arthur's Pass area may more likely be a result of homoplasy rather than a reflection of real world dispersal patterns. Also, while there is no pattern of increasing genetic difference with increasing geographic distance at the widest spatial scale, the fact that the majority of sites between regions have a clear and significant level of pair-wise genetic differentiation indicates that those few that do not may be an anomaly.

Assignment Testing

Assignment testing using the Bayesian statistics based program STRUCTURE (Pritchard et al, 2006) is a commonly used tool to assess the probability that an individual is a migrant or recent descendant (up to three generations) of a migrant. Because of the clear pattern of isolation by distance shown within the Arthur's Pass area I expected that the majority of individuals would be assigned primarily to the stream where they were collected or a neighbouring stream. This is because in a population structured by isolation by distance the majority of dispersers are likely to be from the closest neighbouring population (Slatkin et al, 1998). In the Arthur's Pass area there were two streams (Castle Hill and Cheeseman) that both had a greater than 0.85 proportion of genotypes assigned to the population they were collected in, suggesting very little recent immigration. Furthermore, three other populations (Pylon Gully, Peacock and Moss streams) had the majority of genotypes assigned to either the stream they were collected in or their most geographically proximate stream. The last two Arthurs Pass streams (Waimakariri Bend and Craigieburn) also had a large proportion of genotypes assigned to the stream they were collected in or in the case of Craigieburn a neighbouring stream. These results would appear to support the previously discussed pattern of isolation by distance in this area, and that migration between neighbouring streams is the most prevalent mode of dispersal for this species. However, these two sites also have a large proportion of genotypes assigned to one another (0.416) though they are located 15.7 km apart. This would appear to be evidence of a relatively common long distance dispersal that is at odds with the overall pattern of isolation by distance found in the Arthur's Pass area. Furthermore this conclusion is backed up by the lower pair-wise

genetic structure between these two site compared with pairs of sites located are similar distances from one another in this area (See Fig 3.2).

Because mayflies are comparatively weak flyers (e.g. compared to caddisflies) with a limited adult life span of days it has been suggested that they probably rely on random dispersal mechanisms such as wind (Corkum, 1987). While prevailing winds may provide a general direction for mayfly movement, it has also been suggested that a population that is predominantly dispersed by wind may have lowered success in finding suitable habitat due the inability of the disperser to control the direction of their movement (Kovacs et al, 1996). The overall pattern of isolation by distance in the Arthur's Pass region suggests that at local scales conditions are such that dispersal is more frequent between adjacent streams and that adults are either remaining in their natal area or dispersing laterally only as far as the closest proximate stream. However, rare dispersal events that take adults further than their neighbouring streams do occur.

Assignment testing generated a similar pattern to those in the Φ_{PT} analysis. While the Northland population was assigned almost exclusively to its own region (0.90), there were a few assignments in the Banks Peninsula area that would appear to conflict with the general pattern of dispersal found in the rest of the populations included in this study. There were a number of individuals that STRUCTURE assigned as recent migrants from streams separated by very large geographical distances. The program assigned 10 individuals from Pylon Gully in the Cass area, adjacent to Arthur's Pass National Park to the Kaituna Valley stream located 121 km away on Banks Peninsula. Because of the high levels of genetic differentiation found between sites separated by small geographic distances (See Table 3.2.), and the otherwise fairly local assignments by STRUCTURE (Table 3.3), this would appear to be an anomaly. While

it may be that this is the product of a rare long distance dispersal event, the fact that these two sites have the lowest expected heterozygosity (See Table 3.1) may be causing STRUCTURE to incorrectly assign immigrants from Pylon Gully to Kaituna Valley. The low proportion of polymorphic loci found in the genotypes of these two sites (See Table 3.1) may mean that the Assignment test does not have enough allelic variation to correctly assign the genotypes to the correct site of origin. The high levels of genetic differentiation found between populations in this study, combined with the low levels of expected heterozygosity in these two sites in particular, may indicate that in this case the program has miss-assigned these 10 individuals.

Landscape Genetics

While there may be a host of landscape factors influencing mayfly dispersal, I chose the four variables that I hypothesised would have the greatest affect. Of these landscape factors, only forest cover had a significant relationship with the observed genetic structure of my populations in the Arthur's Pass region. While the strong pattern of isolation by distance found in this area suggests rare but regular step-wise dispersal through the region, it appears that there is considerably more gene flow between streams connected by forest than streams that are separated by open area, even when geographic distance is factored into the model. Therefore, I suggest that forest provides dispersal pathways between streams for *C. humeralis*, enabling a larger more genetically diverse meta-population to exist. This is at odds with the finding that stream insects, in this case a New Zealand caddisfly, do not move laterally into forest a great distance (Collier and Smith, 1997). However, the extent to which stream insects disperse into the surrounding forest appears to vary

taxonomically. For example, Svenson (1974) arranged caddisflies into five groups based on their dispersal distance and direction, with one group appearing to disperse into wooded areas much more than they did upstream. Another hypothesis for the increased genetic connectivity between streams found in forested areas may be that these mayflies are using streams as corridors through forested areas and there is more dispersal happening between these streams because adults are attracted to the suitable habitat found within them. However, there was very little stream connectance found within the forested region in my study area, with only the most likely unsuitable Waimakariri River connecting the three streams. When combined with the fact that this area have a similar level of stream connectance as the area in which stream are separated by open unforested areas, this may suggest that there is more overland between stream dispersal occurring in forested areas than in open areas, at least in this region for this species. This finding corroborates the results of a North American study that found that there was a decrease in the genetic diversity of the forest dwelling mayfly *Ephemerella invaria* in the areas surrounding headwater streams that had a higher percentage of deforestation (Alexander, et al. 2011). While forest cover appears to affect the genetic population structure of more than one species of mayfly, Young et al, (2012) found that in the Australian mayfly, *Ulmerophlebia* sp., there was no genetic difference between fully forested sub-catchments and those with reduced forest cover. However, using more information rich landscape genetics techniques such as circuit theory that rely on GIS mapping, can provide information not only on the comparative extent of forest cover between regions but also their position and their role as potential dispersal corridors between streams (McRae et al, 2005).

Deforestation may be viewed as a surrogate for many co-occurring environmental impacts that lead to larval habitat loss, stream degradation and reduced water quality, and fragmentation (Walsh et al, 2007). Reduction in forest cover will often mean that various stream attributes necessary for mayfly biodiversity will be much reduced and streams flowing through open land are often characterised by increased water temperature, reduced in-stream habitat and oviposition sites, reduced allochthonous inputs, and altered water chemistry (Allan, 2004). Human mediated influences have resulted in many streams being uninhabitable for many of the more sensitive taxa; for example, the Canterbury Plains has few streams that have the ability to support *Coloburiscus humeralis* (Tait et al., 2006). While Harding et al, (2006) found that on Banks Peninsula *Coloburiscus* was either rare or absent from agricultural streams. Furthermore, the results from the landscape modelling back up the genetic data strengthening the argument that dispersing *Coloburiscus* are being affected by a reduction in "stepping stones" to enable dispersal across the larger meta-community (MacArthur and Wilson, 1967). As suitable habitat patches are destroyed, the further a disperser will have to cross to make it to the next intact patch, until the distance becomes so great as to mean that no dispersers are able to make it to the next patch. The consequences for a population of being cut off from the rest of a meta-population have been shown to lead, in many cases, to a higher rate of inbreeding and a higher probability that the population may go extinct (Sacchiri et al, 1999).

It is thought that prior to human colonisation of New Zealand, 800 years before present; much of the land at lower elevations was dominated by forest (McGlone, 1989). The first human colonists, the Maori, were responsible for the first wave of widespread forest destruction. The Maori burned large swathes of forest to plant

bracken fern (*Pteridium aquilinum*), which they used as a food source, to make cross country travel easier, and as a strategy for hunting Moa (Stevens et al., 1988). Late-Holocene charcoal and pollen records indicate that Maori burning of forest was deliberate and systematic (Wethy et al., 2009). This, combined with accidental burning and a drier climate, led to a 68% loss in forest cover at the time of the arrival of Europeans (McGlone, 1989). This second wave of human colonization of the New Zealand archipelago has led to further deforestation. In the 1870s, a growing population, improving roads and the construction of a rail system rapidly increased the rate of forest loss (Arnold, 1994). Deforestation continued well into the 20th century and after the Second World War an increasing amount of high country forest was converted to pasture, or fast growing exotic tree plantations (Leathwick et al., 2003). Human mediated deforestation has led to a reduction in forest cover from an estimated 82% to 23% of the country (Leathwick et al., 2003).

Flying aquatic insects such as mayflies have an adult stage that uses the surrounding terrestrial environment. Adult *Coloburiscus humeralis* have been observed using trees as swarm markers and they may rely on intact riparian vegetation to safely complete their emergence to adulthood (Wethy, 1965). Furthermore, adult stream insects adapted to forested landscapes may not have the physiological capabilities to deal with modified agricultural fields. For example, Carey (2002) found that maximum temperature found in exposed pasture adjacent to the study stream exceeded the thermal tolerance of the adult Leptophlebiidae mayfly. Dispersal in general can be a perilous affair and adult stream insects crossing open agricultural may be more easily visible to keen avian eyes.

For weak flying species such as *Coloburiscus humeralis,* the level of wind may play an important role in dispersal, and may determine whether dispersal is directed or undirected. In open areas it is conceivable that adult aquatic insects may be more reliant on the wind currents to carry them to a suitable habitat, while in forested areas dispersal may be more controlled. Wind-born dispersal by its nature is more likely to be a more random and perilous affair, with a low probability of adults being dispersed to a stream with a viable population of conspecifics. In comparison, animals dispersing through forested areas may rely on powered flight to find an adjacent stream, which because of its proximity and status as forested stream may have a higher chance of containing a population of conspecifics. Furthermore, as the number of viable streams embedded in an open agricultural landscape decreases due to increased anthropogenic impacts, it may be that the probability of a dispersing mayfly finding a suitable stream via wind-born dispersal will decline.

Circuit modelling found that no level of stream connectance was able to influence the population genetic structure that was found among the Arthur's Pass populations. Stream connectance was investigated in previous trapping studies in this eco-region, and many more adult mayflies have been collected flying along stream corridors than flying away from the stream (Winterbourn et al, 2007). The much greater influence of forest connectance and even geographical distance on the gene flow across this area may be explained by several factors. All streams I sampled in the Arthurs Pass region are tributaries of the Waimakariri River and therefore all are connected at the catchment scale. Because of the high levels of genetic structure found between these populations it was therefore unlikely that stream connectance was influencing this structure. However, dendritic tributary network throughout the river is fragmented by open high country pastures with much reduced or entirely absent riparian

vegetation, and it may be that a more functional riparian buffer would provide better pathways for adult movement along the stream corridor.

When higher altitude areas (i.e., 1200m a.s.l.) were included as barriers to dispersal within the resistance model analysis, no relationship was found with the genetic data. However, none of my streams were separated by mountains that might necessitate a dispersing adult to cross a mountain ridge. There is one valley in the Craigieburn area which I expected might provide a dispersal bottleneck because of its relative narrow pass, and therefore mean that some effect of altitude as dispersal barriers might have been found. The last factor, human barriers (specifically: roads, bridges and culverts), also were shown to not act as a barrier to the dispersal of this mayfly. This does not support earlier findings in constricted urban streams that reported a decrease in stream insect diversity upstream of urban culverts (Blakely et al, 2006). However, circuit theory may fail to resolve dispersal patterns affected by roads, as in a large ASCII map, roads may account for a very small number of pixels. Therefore, comparative biodiversity studies like Blakely et al, (2006) may be better equipped to assess whether there is a similar pattern of culvert mediated biodiversity loss in these high country streams as was observed in urban watersheds.

Conclusions

Summary

This study has shown that very limited levels of dispersal appear to be occurring among my populations of *Coloburiscus humeralis*. Even at very small distances, highly significant levels of genetic differentiation were present between sub-populations. However despite these high levels of dispersal limitation evidenced by these results, in the Arthur's Pass region at least, there was a very strong pattern of isolation by distance, meaning that despite significant dispersal barriers present in the landscape, dispersal is occurring in sufficient quantity to produce such a pattern. While genetic diversity was reasonably high in this region, there were three streams in particular that had a comparatively much lower level of genetic diversity. The low levels of dispersal occurring between streams coupled with the low genetic diversity in some of the more isolated populations, may be a sign that the health of the larger meta-population has been affected by the anthropogenic changes to the landscape since human arrival. This is supported by the very strong relationship between the amount of forest cover between streams and the frequency of adult mayflies of this species successfully dispersing between them. Adult *Coloburiscus humeralis* appear to be much more likely to disperse between streams connected by forest than between streams separated by open areas.

The extent of forest cover between streams would appear to support the health of the meta-population as a whole. However it is important to note that because forest cover is so highly correlated with mayfly diversity in Canterbury streams (Harding and Winterbourn, 1995; Harding et al, 2006), whether *Coloburiscus humeralis* preferentially use forests as dispersal corridors or if forest cover is acting as a proxy

for the general health of streams is still to be addressed. It may simply be that as forested streams become more distant from one another dispersal between sites becomes less likely, and areas that have a higher proportion of forest are likely to have more streams with a diverse assemblage of mayfly species (Harding and Winterbourn, 1995). Either way, there seems to be good evidence that intact native forest promotes healthy meta-populations, with high genetic diversity and more frequently interacting sub-populations.

Practical Implications

This study supports the large body of evidence that suggests that by maintaining forests surrounding streams; benthic invertebrate assemblages benefit (Harding and Winterbourn, 1995; Harding et al, 2006; Death and Collier, 2010; Alexander et al, 2011). Furthermore, the results of this study show that human mediated changes to the landscape, in this case deforestation, can have complex negative effects on the species that reside there (Dirzo and Raven, 2003). This provides further evidence that the terrestrial and freshwater ecosystems that flow through them are intricately linked (Abell et al, 2002). Therefore if other taxa of New Zealand stream invertebrates are similarly dispersal limited, the maintenance of existing forested areas may be crucial for preserving the current biodiversity of New Zealand freshwater invertebrates.

Due to the limited dispersal occurring within this species at very small distances, it appears that *Coloburiscus humeralis* does not regularly disperse over long distances. Thus, these results suggest that stream remediation efforts are most likely to maximise recolonisation by these species if carried out in streams that are as close as

possible to source populations. If stream remediation projects are enacted in catchments that are distant from such source populations, managers may consider translocating larvae to these newly created habitats. The findings of this study support the increased focus on riparian remediation that has occurred in New Zealand in recent years (Greenwood et al, 2012). Streams form dendritic pathways across the landscape, and by replacing riparian buffers along waterways running through agricultural areas, not only will the in stream environment be enhanced (Wenger, 1999), it may also promote greater inter sub-population dispersal; and therefore a more robust and genetically diverse meta-population.

Future Work

While this study provides good evidence that forest cover effects the dispersal of *Coloburiscus humeralis* in the Arthur's Pass area, extending this study across multiple region and multiple areas would be beneficial. This would help to ensure that the dispersal patterns found in this study and the factors found to affect them are able to be generalised across New Zealand for this endemic mayfly species. Also, previous studies on mayfly dispersal have found significant genetic differentiation between populations living in different reaches within the same stream (Morris, 2005), and replicating this part of the Morris study (2005) with a fast mutating marker such as AFLPs would help to confirm this. Despite some early technical difficulties with the AFLP process I was able generate a data set that answered the majority of the questions that I had previous to the study. However, the development of co-dominant markers, like microsatellites, for this species would enable a more precise measurement of the level of inbreeding and the migration rate.

This study has shown that landscape genetic techniques can be successfully applied to the study of freshwater invertebrate populations and using such techniques to uncover both dispersal corridors and barriers in a variety of taxa will help to develop successful management plans for freshwater ecosystems. While there have been a few population genetic studies focusing on flying freshwater insects in New Zealand (Hogg et al, 2002), further work using rapidly mutating markers across a range of taxa would help to develop a better picture of the factors influencing biodiversity in New Zealand streams.

References

Abell, R., Thieme, M., Dinerstein, E., Olson, D., (2002). Conservation biology for the biodiversity crisis: a freshwater follow-up. Conservation Biology 16, 1435–1437.

Aerts, R., & Berendse, F. (1988). The effect of increased nutrient availability on vegetation dynamics in wet heathlands. *Vegetatio*, *76*(1-2), 63-69.

Alexander, L. C., Hawthorne, D. J., Palmer, M. A., & Lamp, W. O. (2011). Loss of genetic diversity in the North American mayfly Ephemerella invaria associated with deforestation of headwater streams. *Freshwater Biology*,*56*(7), 1456-1467.

Arnaud-Haond, S., Alberto, F., Teixeira, S., Procaccini, G., Serrao, E. A., & Duarte, C. M. (2005). Assessing genetic diversity in clonal organisms: low diversity or low resolution? Combining power and cost efficiency in selecting markers. *Journal of Heredity*, *96*(4), 434-440.

Barnosky, A. D., Matzke, N., Tomiya, S., Wogan, G. O., Swartz, B., Quental, T. B., ... & Ferrer, E. A. (2011). Has the Earth/'s sixth mass extinction already arrived?. *Nature*, *471*(7336), 51-57.

Benjamini, Y. and Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. Journal of the Royal Statistical Society 57, 289–300.

Berlocher, S. H., & Feder, J. L. (2002). Sympatric speciation in phytophagous insects: moving beyond controversy?. *Annual review of entomology*, *47*(1), 773-815.

Bilton, D. T., Freeland, J. R., & Okamura, B. (2001). Dispersal in freshwater invertebrates. *Annual review of ecology and systematics*, 159-181.

Blakely, T. J., Harding, J. S., McIntosh, A. R., & Winterbourn, M. J. (2006). Barriers to the recovery of aquatic insect communities in urban streams. *Freshwater Biology*, *51*(9), 1634-1645.

Bohonak, A. J., & Jenkins, D. G. (2003). Ecological and evolutionary significance of dispersal by freshwater invertebrates. *Ecology letters*, *6*(8), 783-796.

Bond, N. R., & Downes, B. J. (2003). The independent and interactive effects of fine sediment and flow on benthic invertebrate communities characteristic of small upland streams. *Freshwater Biology*, *48*(3), 455-465.

Bonin A, Bellemain E, Bronken Eidesen P. (2004). How to track and assess genotyping errors in population genetics studies. Molecular Ecology 13, 3261–3273.

Bossart, J. L., & Pashley Prowell, D. (1998). Genetic estimates of population structure and gene flow: limitations, lessons and new directions. *Trends in Ecology & Evolution*, *13*(5), 202-206.

Boulton, A. J., & Lake, P. S. (2008). Effects of drought on stream insects and its ecological consequences. *Aquatic insects: challenges to populations, CABI Publishing, Wallingford*, 81-102.

Bowler, D. E., & Benton, T. G. (2005). Causes and consequences of animal dispersal strategies: relating individual behaviour to spatial dynamics. *Biological Reviews*, *80*(2), 205-225.

Bremer, B., & Eriksson, O. (1992). Evolution of fruit characters and dispersal modes in the tropical family Rubiaceae. *Biological Journal of the Linnean Society*, *47*(1), 79-95.

Briers, R. A., Cariss, H. M., & Gee, J. H. (2002). Dispersal of adult stoneflies (Plecoptera) from upland streams draining catchments with contrasting land-use. *Archiv für Hydrobiologie*, *155*(4), 627-644.

Briers, R. A., Gee, J. H., Cariss, H. M., & Geoghegan, R. (2004). Inter-population dispersal by adult stoneflies detected by stable isotope enrichment. *Freshwater Biology*, *49*(4), 425-431.

Brittain, J. E. (1982). Biology of mayflies. *Annual Review of Entomology*,*27*(1), 119-147.

Brittain, J. E. (1989). Life history strategies in Ephemeroptera and Plecoptera. In *Mayflies and stoneflies: life histories and biology* (pp. 1-12). Springer Netherlands.

Buckley, T. R., Marske, K., & Attanayake, D. (2010). Phylogeography and ecological niche modelling of the New Zealand stick insect Clitarchus hookeri (White) support survival in multiple coastal refugia. *Journal of Biogeography*, *37*(4), 682-695.

Butchart, S. H., Walpole, M., Collen, B., van Strien, A., Scharlemann, J. P., Almond, R. E., ... & Watson, R. (2010). Global biodiversity: indicators of recent declines. *Science*, *328*(5982), 1164-1168.

Caballero, A., Quesada, H., & Rolán-Alvarez, E. (2008). Impact of amplified fragment length polymorphism size homoplasy on the estimation of population genetic diversity and the detection of selective loci. *Genetics*, *179*(1), 539-554.

Cadet, C., Ferrière, R., Metz, J. A., & van Baalen, M. (2003). The evolution of dispersal under demographic stochasticity. *The American Naturalist*, *162*(4), 427-441.

Calderón-Cortés, N; Quesada, M; Cano-Camacho, H; Zavala-Páramo, G. A. (2010). A simple and rapid method for DNA isolation from xylophagous insects. *International Journal of Molecular Science 11*, 5056–5064.

Carpenter, F. M., & Burnham, L. (1985). The geological record of insects.*Annual Review of Earth and Planetary Sciences*, *13*, 297

Caudill, C. C. (2003). Measuring dispersal in a metapopulation using stable isotope enrichment: high rates of sex-biased dispersal between patches in a mayfly metapopulation. *Oikos*, *101*(3), 624-630.

Charlesworth, D., & Charlesworth, B. (1987). Inbreeding depression and its evolutionary consequences. *Annual review of ecology and systematics*, *18*, 237-268.

Cicero, C. (2004). Barriers to sympatry between avian sibling (PARIDAE: BAEOLOPHUS) In local secondary contact. *Evolution*, *58*(7), 1573-1587.

Clavel, J., Julliard, R., & Devictor, V. (2010). Worldwide decline of specialist species: toward a global functional homogenization?. *Frontiers in Ecology and the Environment*, *9*(4), 222-228.

Clobert, J., Galliard, L., Cote, J., Meylan, S., & Massot, M. (2009). Informed dispersal, heterogeneity in animal dispersal syndromes and the dynamics of spatially structured populations. *Ecology letters*, *12*(3), 197-209.

Coffin, A. W. (2007). From roadkill to road ecology: a review of the ecological effects of roads. *Journal of Transport Geography*, *15*(5), 396-406.

Collier, K. J., & Quinn, J. M. (2003). Land-use influences macroinvertebrate community response following a pulse disturbance. *Freshwater Biology*,*48*(8), 1462-1481.

Collier, K. J., & Smith, B. J. (1995). Sticky trapping of adult mayflies, stoneflies and caddisflies alongside three contrasting streams near Hamilton, New Zealand. *New Zealand natural sciences*, *22*, 1-9.

Collier, K. J., & Smith, B. J. (1997). Dispersal of adult caddisflies (Trichoptera) into forests alongside three New Zealand streams.*Hydrobiologia*, *361*(1-3), 53-65.

Cote, J., & Clobert, J. (2007). Social personalities influence natal dispersal in a lizard. *Proceedings of the Royal Society B: Biological Sciences*, *274*(1608), 383-390.

Cushman, S. A., McKelvey, K. S., Hayden, J., & Schwartz, M. K. (2006). Gene flow in complex landscapes: testing multiple hypotheses with causal modeling. *The American Naturalist*, *168*(4), 486-499.

Cushman, S; Wasserman, T; Landguth, E; Shirk, A. (2013). Re-Evaluating Causal Modeling with Mantel Tests in Landscape Genetics. Diversity 5, 51–72.

Darwin, C. (1859). On the origins of species by means of natural selection.*London: Murray*.

Death R.G. & Collier K.J. (2010) Measuring stream macroinvertebrate responses to gradients of vegetation cover: when is enough enough? Freshwater Biology, 55, 1447–1464.

Denno, R. F., Roderick, G. K., Peterson, M. A., Huberty, A. F., Dobel, H. G., Eubanks, M. D., ... & Langellotto, G. A. (1996). Habitat persistence underlies intraspecific

variation in the dispersal strategies of planthoppers. *Ecological Monographs*, *66*(4), 389-408.

Didham, R. K., Blakely, T. J., Ewers, R. M., Hitchings, T. R., Ward, J. B., & Winterbourn, M. J. (2012). Horizontal and vertical structuring in the dispersal of adult aquatic insects in a fragmented landscape. *Fundamental and Applied Limnology/Archiv für Hydrobiologie*, *180*(1), 27-40.

Didham, R. K., Hammond, P. M., Lawton, J. H., Eggleton, P., & Stork, N. E. (1998). Beetle species responses to tropical forest fragmentation.*Ecological Monographs*, *68*(3), 295-323.

Dirzo R. & Raven P.H. (2003) Global state of biodiversity and loss. Annual Review of Environment and Resources, 28, 137–167.

Downes, B. J., & Reich, P. (2008). What is the spatial structure of stream insect populations? Dispersal behaviour of different life-history stages.*Aquatic insects: challenges to populations*, 184-203.

Doyle, P. and Snell. J. (1984). Random walks and electric networks, volume 22. Mathematical Association America, New York.

Dudley, R. (2002). *The biomechanics of insect flight: form, function, evolution*. Princeton University Press.

Dytham, C. (2009). Evolved dispersal strategies at range margins. *Proceedings of the Royal Society B: Biological Sciences*, *276*(1661), 1407-1413.

Eriksson, O., & Jakobsson, A. (1999). Recruitment trade-offs and the evolution of dispersal mechanisms in plants. *Evolutionary Ecology*, *13*(4), 411-423.

Ewers, R. M., & Didham, R. K. (2006). Confounding factors in the detection of species responses to habitat fragmentation. *Biological Reviews*, *81*(1), 117-142.

Excoffier L., Smouse P.E. and Quattro J.M. (1992). Analysis of molecular variance inferred from metric distances among DNA haplotypes – Application to Human Mitochondrial – DNA Restriction Data. Genetics 131: 479–491.

Falush, D; Stephens, M; Pritchard, J. (2007). Inference ofpopulation structure using multilocus genotype data: dominant markers and null alleles. Molecular Ecology Notes 7, 574–578.

Frankham, R. (2010). Challenges and opportunities of genetic approaches to biological conservation. *Biological conservation*, *143*(9), 1919-1927.

Fung, F., Watts, G., Lopez, A., Orr, H. G., New, M., & Extence, C. (2013). Using large climate ensembles to plan for the hydrological impact of climate change in the freshwater environment. *Water resources management*, 1-22.

Gadgil, M. (1971). Dispersal: population consequences and evolution. *Ecology*,*52*(2), 253-261.

Gibbins, C., Batalla, R. J., & Vericat, D. (2010). Invertebrate drift and benthic exhaustion during disturbance: Response of mayflies (Ephemeroptera) to increasing shear stress and river-bed instability. *River Research and Applications*, *26*(4), 499-511.

Gibbs, H. L., Gibbs, K. E., Siebenmann, M., & Collins, L. (1998). Genetic differentiation among populations of the rare mayfly Siphlonisca aerodromia Needham. *Journal of the North American Benthological Society*, 464-474.

Goslee, S.C; and Urban, D.L. (2007). The ecodist package for dissimilarity-based analysis of ecological data. Journal of Statistical Software 22(7), 1-19.

Greenwood, M. J., Harding, J. S., Niyogi, D. K., & McIntosh, A. R. (2012). Improving the effectiveness of riparian management for aquatic invertebrates in a degraded agricultural landscape: stream size and land-use legacies. *Journal of Applied Ecology*, *49*(1), 213-222.

Gu, H. N., Hughes, J. & Dorn, S. (2006). Trade-off between mobility and fitness in Cydia pomonella L. (Lepidoptera: Tortricidae). Ecological Entomology, 31, 68e74.

Guichoux, E., Lagache, L., Wagner, S., Chaumeil, P., Léger, P., Lepais, O., & Petit, R. J. (2011). Current trends in microsatellite genotyping. *Molecular Ecology Resources*, *11*(4), 591-611.

Hamilton, W.D. (1964) The genetical evolution of social behaviour. I & II. J. Theor. Biol. 7, 1–52.

Hanski, I. (1999). Habitat connectivity, habitat continuity, and metapopulations in dynamic landscapes. *Oikos*, 209-219.

Harding, J. S., & Winterbourn, M. J. (1993). Life history and production of Coloburiscus humeralis (Ephemeroptera: Oligoneuriidae) in two South Island high-country streams, New Zealand. *New Zealand journal of marine and freshwater research*, *27*(4), 445-451.

Harding, J. S., & Winterbourn, M. J. (1995). Effects of contrasting land use on physico-chemical conditions and benthic assemblages of streams in a Canterbury (South Island, New Zealand) river system. *New Zealand journal of marine and freshwater research*, *29*(4), 479-492.

Harding, J. S., Claassen, K., & Evers, N. (2006). Can forest fragments reset physical and water quality conditions in agricultural catchments and act as refugia for forest stream invertebrates?. *Hydrobiologia*, *568*(1), 391-402.

Harker, J. E. (1992). Swarm behaviour and mate competition in mayflies (Ephemeroptera). *Journal of Zoology*, *228*(4), 571-587.

Hastings, A. (1983). Can spatial variation alone lead to selection for dispersal?.*Theoretical Population Biology*, *24*(3), 244-251.

Hättenschwiler, S., & Vitousek, P. M. (2000). The role of polyphenols in terrestrial ecosystem nutrient cycling. *Trends in Ecology & Evolution*, *15*(6), 238-243.

Hedrick, P. W. (2005). A standardized genetic differentiation measure. Evolution 59. 1633-1638.

Hinton, H. E. (1948).On the origin of the function of the pupal. *Transactions of the Royal Entomological Society of London*, *99*(12), 395-409.

Hoekstra, J. M., Boucher, T. M., Ricketts, T. H., & Roberts, C. (2005). Confronting a biome crisis: global disparities of habitat loss and protection.*Ecology Letters*, *8*(1), 23-29.

Hogg, I D., Willmann-Huerner p., and Stevens, M. (2002) Population genetic structures of two New Zealand stream insects: Archichauliodes diversus (Megaloptera) and Coloburiscus humeralis (Ephemeroptera). *New Zealand Journal of Marine and Freshwater Research,* 36 (3) 491-501.

Holt, R. D. (1985). Population dynamics in two-patch environments: some anomalous consequences of an optimal habitat distribution. *Theoretical population biology*, *28*(2), 181-208.

Hoskin, C. J., Higgie, M., McDonald, K. R., & Moritz, C. (2005). Reinforcement drives rapid allopatric speciation. *Nature*, *437*(7063), 1353-1356.

Howe, H. F., & Smallwood, J. (1982). Ecology of seed dispersal. *Annual review of ecology and systematics*, *13*, 201-228.

Treier, U (2010). AFLP Protocol. http://urstreier.net/uploads/pdf/Treier-2010-AFLP_Protocol.pdf.

Hubbell, S. P. (2001). The unified neutral theory of biodiversity and biogeography *(MPB-32)* (Vol. 32). Princeton University Press.

Hufford, K. M., Krauss, S. L., & Veneklaas, E. J. (2012). Inbreeding and outbreeding depression in Stylidium hispidum: implications for mixing seed sources for ecological restoration. *Ecology and evolution*, *2*(9), 2262-2273.

Hughes, J. M., Hillyer, M., & Bunn, S. E. (2003). Small-scale patterns of genetic variation in the mayfly Bungona narilla (Ephemeroptera: Baetidae) in rainforest streams, south-east Queensland. *Freshwater Biology*, *48*(4), 709-717.

Hughes, J. M., Schmidt, D. J., & Finn, D. S. (2009). Genes in streams: using DNA to understand the movement of freshwater fauna and their riverine habitat. *BioScience*, *59*(7), 573-583.

Imbert, E., & Ronce, O. (2001). Phenotypic plasticity for dispersal ability in the seed heteromorphic *Crepis sancta* (Asteraceae). *Oikos*, *93*(1), 126-134.

Ims, R. A., & Andreassen, H. P. (2005). Density-dependent dispersal and spatial population dynamics. *Proceedings of the Royal Society B: Biological Sciences*,*272*(1566), 913-918.

Intergovernmental Panel on Climate Change (2001). Third Assessment Report of the Intergovernmental Panel on Climate Change IPCC (WG I & II). Cambridge Univ. Press, Cambridge.

Jackson, J. K., & Resh, V. H. (1989). Distribution and abundance of adult aquatic insects in the forest adjacent to a northern California stream. *Environmental entomology*, *18*(2), 278-283.

Jowett, I. G., Richardson, J., & Boubee, J. A. T. (2009). Effects of riparian manipulation on stream communities in small streams: two case studies. *New Zealand Journal of Marine and Freshwater Research*, *43*(3), 763-774.

Keyghobadi, N., Roland, J. E. N. S., & Strobeck, C. (2005). Genetic differentiation and gene flow among populations of the alpine butterfly, *Parnassius smintheus*, vary with landscape connectivity. *Molecular Ecology*, *14*(7), 1897-1909.

Klug, W. S., & Cummings, M. R. (2003). *Concepts of genetics* (No. Ed. 7). Pearson Education, Inc.

Kovats, Z., Ciborowski, J. A. N., & Corkum, L. (1996). Inland dispersal of adult aquatic insects. *Freshwater biology*, *36*(2), 265-276.

Krauss, J., Bommarco, R., Guardiola, M., Heikkinen, R. K., Helm, A., Kuussaari, M., & Steffan-Dewenter, I. (2010). Habitat fragmentation causes immediate and time-delayed biodiversity loss at different trophic levels. *Ecology letters*, *13*(5), 597-605.

Krauss, S. L. (2000). Accurate gene diversity estimates from amplified fragment length polymorphism (AFLP) markers. *Molecular Ecology*, *9*(9), 1241-1245.

Lachlan, R. F., & Servedio, M. R. (2004). Song learning accelerates allopatric speciation. *Evolution*, *58*(9), 2049-2063.

Lancaster, J., Downes, B. J., & Arnold, A. (2011). Lasting effects of maternal behaviour on the distribution of a dispersive stream insect. *Journal of Animal Ecology*, *80*(5), 1061-1069.

Legendre, P. (2000). Comparison of permutation methods for the partial correlation and partial mantel tests. Journal of Statistics and Computational Simulations 67, 37-73.

Legendre, P.; and Fortin, M.J. (2010). Comparison of the Mantel test and alternative approaches for detecting complex multivariate relationships in the spatial analysis of genetic data. Molecular Ecology Resources 10, 831–844.

Leturque, H., & Rousset, F. (2003). Joint evolution of sex ratio and dispersal: conditions for higher dispersal rates from good habitats. *Evolutionary Ecology*,*17*(1), 67-84.

Levin, S. A., Muller-Landau, H. C., Nathan, R., & Chave, J. (2003). The ecology and evolution of seed dispersal: a theoretical perspective. *Annual Review of Ecology, Evolution, and Systematics*, 575-604.

Lynch, J. (1989). The gauge of speciation: on the frequencies of modes of speciation.

Lynch, M; and Milligan, B.G. (1994). Analysis of population genetic structure with RAPD markers, Molecular Ecology 3, 91–99.

MacArthur, R. H., & Wilson, E. O. (1967). The theory of island biography.*Univ. Press, Princeton, NJ*, 203.

Macneale, K. H., Peckarsky, B. L., & Likens, G. E. (2004). Contradictory results from different methods for measuring direction of insect flight.*Freshwater Biology*, *49*(10), 1260-1268.

Mallet, J. (1995). A species definition for the modern synthesis. *Trends in Ecology & Evolution*, *10*(7), 294-299.575-604.

Malmqvist, B. (2000). How does wing length relate to distribution patterns of stoneflies (Plecoptera) and mayflies (Ephemeroptera)? *Biological Conservation*, *93*(2), 271-276.

Manel, S., Schwartz, M. K., Luikart, G., & Taberlet, P. (2003). Landscape genetics: combining landscape ecology and population genetics. *Trends in Ecology & Evolution*, *18*(4), 189-197.

Mantel, N. and Valand, R. S. (1970). A technique of nonparametric multivariate analysis. Biometrics 26, 547-558.

Markert, J. A., Champlin, D. M., Gutjahr-Gobell, R., Grear, J. S., Kuhn, A., McGreevy, T. J., & Nacci, D. E. (2010). Population genetic diversity and fitness in multiple environments. *BMC evolutionary biology*, *10*(1), 205.

McGlone, M.S., (1989). The Polynesian settlement of New Zealand in relation to environmental and biotic changes. New Zealand Journal of Ecology 12, 115–129.

McLay, C. (1970). A theory concerning the distance travelled by animals entering the drift of a stream. *Journal of the Fisheries Board of Canada*,*27*(2), 359-370.

McPeek, M. A., & Holt, R. D. (1992). The evolution of dispersal in spatially and temporally varying environments. *American Naturalist*, 1010-1027.

McRae, B. H., & Beier, P. (2007). Circuit theory predicts gene flow in plant and animal populations. *Proceedings of the National Academy of Sciences*,*104*(50), 19885-19890.

McRae, B. H., Dickson, B. G., Keitt, T. H., & Shah, V. B. (2008). Using circuit theory to model connectivity in ecology, evolution, and conservation.*Ecology*, *89*(10), 2712-2724.

McWethy, D. B., Whitlock, C., Wilmshurst, J. M., McGlone, M. S., & Li, X. (2009). Rapid deforestation of south island, New Zealand, by early Polynesian fires. *The Holocene*, *19*(6), 883-897.

Meirmans, P.G. (2006). Using the AMOVA Framework to Estimate a Standardized Genetic Differentiation Measure. Evolution 60 (11). 2399-2402.

Meudt, H. M., & Clarke, A. C. (2007). Almost forgotten or latest practice? AFLP applications, analyses and advances. *Trends in plant science*, *12*(3), 106-117.

Mickett, K., Morton, C., Feng, J., Li, P., Simmons, M., Cao, D., ... & Liu, Z. (2003). Assessing genetic diversity of domestic populations of channel catfish (Ictalurus punctatus) in Alabama using AFLP markers. *Aquaculture*, *228*(1), 91-105.

Moore, J. C., Loggenberg, A., & Greeff, J. M. (2006). Kin competition promotes dispersal in a male pollinating fig wasp. *Biology letters*, *2*(1), 17-19.

Morris, P. (2005). Genetic population structure of four taxa of aquatic insect at three hierarchical spatial scales (Master's Thesis). University of Canterbury.

Moyle, L. C. (2006). Correlates of genetic differentiation and isolation by distance in 17 congeneric Silene species. *Molecular Ecology*, *15*(4), 1067-1081.

Muller, K. (1982). The colonisation cycle of freshwater insects. Oecologia 52: 202–207.

Nathan, R. (2006). Long-distance dispersal of plants. *Science*, *313*(5788), 786-788.

Neigel, J.E. (2002). Is F-ST obsolete? Conservation Genetics 3, 167–173.

Nelson-Flower, M. J., Hockey, P. A., O'Ryan, C., Raihani, N. J., du Plessis, M. A., & Ridley, A. R. (2011). Monogamous dominant pairs monopolize reproduction in the cooperatively breeding pied babbler. *Behavioral Ecology*, *22*(3), 559-565.

Neves, R. J. (1979). Movements of larval and adult Pycnopsyche guttifer (Walker)(Trichoptera: Limnephilidae) along Factory Brook, Massachusetts.*American Midland Naturalist*, 51-58.

Parvinen, K., Dieckmann, U., Gyllenberg, M., & Metz, J. A. (2003). Evolution of dispersal in metapopulations with local density dependence and demographic stochasticity. *Journal of evolutionary biology*, *16*(1), 143-153.

Peakall, R. and Smouse, P.E. (2012). GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research – an update. Bioinformatics 28, 2537-2539.

Perrin, N., & Mazalov, V. (2000). Local competition, inbreeding, and the evolution of sex-biased dispersal. *The American Naturalist*, *155*(1), 116-127.

Petersen, I., Masters, Z., Hildrew, A. G., & Ormerod, S. J. (2004). Dispersal of adult aquatic insects in catchments of differing land use. *Journal of Applied Ecology*, *41*(5), 934-950.

Poethke, H. J., Pfenning, B., & Hovestadt, T. (2010). The relative contribution of individual and kin selection to the evolution of density-dependent dispersal rates.

Pritchard, J.K. (2000). Inference of population structure using multilocus genotype data. Genetics 155, 945–959.

Pusey, A. E. (1987). Sex-biased dispersal and inbreeding avoidance in birds and mammals. *Trends in Ecology & Evolution*, *2*(10), 295-299.

Rayfield, B; Forti,n M.J; Fall, A. (2010).The sensitivity of least-cost habitat graphs to relative cost surface values. Landscape Ecology, 25, 519–532.

Reed, David H., David A. Briscoe, and Richard Frankham. (2002) Inbreeding and extinction: the effect of environmental stress and lineage. *Conservation Genetics* 3.3: 301-307.

Reineke, A; Karlovsky, P; Zebitz C.P.W. (1998). Preparation and purification of DNA from insects for AFLP analysis. Insect Molecular Biology. 7(1):95–99.

Relethford, J. (2004). Global patterns of isolation by distance based on genetic and morphological data. *Human biology*, *76*(4), 499-513.

Richardson, J. L. (2012). Divergent landscape effects on population connectivity in two co-occurring amphibian species. Molecular Ecology. 21:4437–4451.

Roff, D. A. (1986). The evolution of wing dimorphism in insects. *Evolution*, 1009-1020.

Rogers, S.O. and Bendich, AJ. (1985). Extraction of DNA from milligram amounts of fresh, herbarium and mummified plant tissues. Plant Molecular Biology 5: 69-76.

Ronce, O. (2007). How does it feel to be like a rolling stone? Ten questions about dispersal evolution. *Annu. Rev. Ecol. Evol. Syst.*, *38*, 231-253.

Ronce, O., Brachet, S., Olivieri, I., Gouyon, P. H., & Clobert, J. (2005). Plastic changes in seed dispersal along ecological succession: theoretical predictions from an evolutionary model. *Journal of Ecology*, *93*(2), 431-440.

Rousset, F., & Gandon, S. (2002). Evolution of the distribution of dispersal distance under distance-dependent cost of dispersal. *Journal of Evolutionary Biology*, *15*(4), 515-523.

Rubtsova, G. A., Afanasiev, K. I., Malinina, T. V., Shitova, M. V., Rakitskaya, T. A., Prokhorovskaya, V. D., & Zhivotovsky, L. A. (2008). Differentiation of chum salmon Oncorhynchus keta Wallbaum populations as revealed with microsatellite and allozyme markers: A comparative study. *Russian Journal of Genetics*, *44*(7), 841-848.

Ryan, P. A. (1991). Environmental effects of sediment on New Zealand streams: a review. *New Zealand journal of marine and freshwater research*, *25*(2), 207-221.

Saccheri, I., Kuussaari, M., Kankare, M., Vikman, P., Fortelius, W., & Hanski, I. (1998). Inbreeding and extinction in a butterfly metapopulation. *Nature*,*392*(6675), 491-494.

Sackett, L.C; Cross, T; Jones, R.T; Johnson. W.C; Ballare. K; Ray, C; Collinge, S.K; Martin, A.P. (2012). Connectivity of prairie dog colonies in an altered landscape: inferences from analysis of microsatellite DNA variation. Conservation Genetics 13, 407–418.

Sax, D. F., & Brown, J. H. (2000). The paradox of invasion. *Global Ecology and Biogeography*, *9*(5), 363-371.

Scheffer, M., Carpenter, S., Foley, J. A., Folke, C., & Walker, B. (2001). Catastrophic shifts in ecosystems. *Nature*, *413*(6856), 591-596.

Schneider, C., Dover, J., & Fry, G. L. (2003). Movement of two grassland butterflies in the same habitat network: the role of adult resources and size of the study area. *Ecological Entomology*, *28*(2), 219-227.

Schwartz, M.K; Copeland, J.P; Anderson, N.J; Squires, J.R; Inman, R.M. (2009). Wolverine gene flow across a narrow climatic niche. Ecology 90 (11), 3222–3232.

Shafer, A.B.A; Northrup, J.M; White, K.S; Boyce, M.S; Côté, S.D. (2012). Habitat selection predicts genetic relatedness in an alpine ungulate. Ecology 93: 1317–1329.

Shirasawa, K., Kishitani, S., & Nishio, T. (2004). Conversion of AFLP markers to sequence-specific markers for closely related lines in rice by use of the rice genome sequence. *Molecular breeding*, *14*(3), 283-292.

Shpak, M. (2005). The role of deleterious mutations in allopatric speciation.*Evolution*, *59*(7), 1389-1399.

Slatkin, M. (1985). Gene flow in natural populations. *Annual review of ecology and systematics*, *16*, 393-430.

Slatkin, M. (1993). Isolation by distance in equilibrium and nonequilibrium populations. Evolution, 47, 264-279.

Sloggett, J. J., & Weisser, W. W. (2002). Parasitoids induce production of the dispersal morph of the pea aphid, Acyrthosiphon pisum. *Oikos*, *98*(2), 323-333.

Smith, P. J. (2009). Genetic principles for freshwater restoration in New Zealand. *New Zealand Journal of Marine and Freshwater Research*, *43*(3), 749-762.

Smith, P. J., & Collier, K. J. (2001). Allozyme diversity and population genetic structure of the caddisfly Orthopsyche fimbriata and the mayfly Acanthophlebia cruentata in New Zealand streams. *Freshwater Biology*,*46*(6), 795-805.

Smith, V. H., & Schindler, D. W. (2009). Eutrophication science: where do we go from here?. *Trends in Ecology & Evolution*, *24*(4), 201-207.

Smith, V. H., Tilman, G. D., & Nekola, J. C. (1999). Eutrophication: impacts of excess nutrient inputs on freshwater, marine, and terrestrial ecosystems.*Environmental pollution*, *100*(1), 179-196.

Smouse, P.E; Long, J.C; Sokal, R.R. (1986). Multiple regression and correlation extensions of the Mantel test of matrix correspondence. Systematic Zoology, 35, 627–632.

Spear, S.F; Storfer, A. (2010). Anthropogenic and natural disturbance lead to differing patterns of gene flow in the Rocky Mountain tailed frog, Ascaphus montanus. Biological Conservation, 143, 778–786.

Stipp, J. J., & McDougall, I. (1968). Geochronology of the Banks Peninsula volcanoes, New Zealand. *New Zealand journal of geology and geophysics*,*11*(5), 1239-1258.

Storfer, A., Murphy, M. A., Evans, J. S., Goldberg, C. S., Robinson, S., Spear, S. F., ... & Waits, L. P. (2006). Putting the 'landscape'in landscape genetics.*Heredity*, *98*(3), 128-142.

Storfer, A; Murphy, M.A; Evans, J.S; Goldberg, C.S; Robinson, S; Spear, S.F; Dezzani, R; Delmelle, E; Vierling, L; Waits, L.P. (2007). Putting the 'landscape' in landscape genetics. Heredity, 98, 128–142.

Stutz, H. L., Shiozawa, D. K., & Evans, R. P. (2010). Inferring dispersal of aquatic invertebrates from genetic variation: a comparative study of an amphipod and mayfly in Great Basin springs. *Journal of the North American Benthological Society*, *29*(3), 1132-1147.

Sunnucks, P. (2000). Efficient genetic markers for population biology. *Trends in Ecology & Evolution*, *15*(5), 199-203.

Svensson, B. W. (1974). Population movements of adult Trichoptera at a South Swedish stream. *Oikos*, 157-175.

Tait, P. R., & Cullen, R. (2006). Some external costs of dairy farming in Canterbury. In *Australian Agricultural and Resource Economics Society Annual Conference* (pp. 8-10).

Thompson, R., & Townsend, C. (2006). A truce with neutral theory: local deterministic factors, species traits and dispersal limitation together determine patterns of diversity in stream invertebrates. *Journal of Animal Ecology*, *75*(2), 476-484.

Tilman, D., Fargione, J., Wolff, B., DAntonio, C., Dobson, A., Howarth, R. et al. (2001). Forecasting agriculturally driven global environmental change. Science, 292, 281–284.

Torgo, L. (2010). *Data Mining using R: learning with case studies*, CRC Press (ISBN: 9781439810187).

Townsend, C. R. (1996). Invasion biology and ecological impacts of brown trout (Salmo trutta) in New Zealand. *Biological Conservation*, *78*(1), 13-22.

Townsend, C. R. (2003). Individual, population, community, and ecosystem consequences of a fish invader in New Zealand streams. *Conservation Biology*, *17*(1), 38-47.

Vekemans , X; Beauwens, T; Lemaire, M; Roldán-Ruiz, I. (2002). Data from amplified fragment length polymorphism (AFLP) markers show indication of size homoplasy and of a relationship between degree of homoplasy and fragment size. *Molecular Ecology* 11(1): 139-151.

Vos, P., Hogers, R., Bleeker, M., Reijans, M., van De Lee, T., Hornes, M., ... & Zabeau, M. (1995). AFLP: a new technique for DNA fingerprinting. *Nucleic acids research*, *23*(21), 4407-4414.

Vos, P; Hogers, R; Bleeker, M; Reijans, M; van De Lee, T; Hornes, M; Zabeau, M. (1995). AFLP: a new technique for DNA fingerprinting. *Nucleic acids research*, *23*(21), 4407-4414.

Wagenhoff, A., Townsend, C. R., & Matthaei, C. D. (2012). Macroinvertebrate responses along broad stressor gradients of deposited fine sediment and dissolved

nutrients: a stream mesocosm experiment. *Journal of Applied Ecology*,*49*(4), 892-902

Walther, G. R., Post, E., Convey, P., Menzel, A., Parmesan, C., Beebee, T. J., & Bairlein, F. (2002). Ecological responses to recent climate change. *Nature*, *416* (6879), 389-395.

Wenger, S. (1999). A review of the scientific literature on riparian buffer width, extent and vegetation.

Weising K, Nybom H, Wolff K, Meyer W (1995) *DNA fingerprinting in plants and fungi*. CRC Press, Boca Raton, USA.

Wiens, J. J., & Donoghue, M. J. (2004). Historical biogeography, ecology and species richness. *Trends in Ecology & Evolution*, *19*(12), 639-644.

Winkworth, R. C., Wagstaff, S. J., Glenny, D., & Lockhart, P. J. (2002). Plant dispersal news from New Zealand. *Trends in Ecology & Evolution*, *17*(11), 514-520.

Winterbourn, M. J., Chadderton, W. L., Entrekin, S. A., Tank, J. L., & Harding, J. S. (2007). Distribution and dispersal of adult stream insects in a heterogeneous montane environment. *Fundamental and Applied Limnology/Archiv für Hydrobiologie*, *168*(2), 127-135.

Wisely, B. (1965). Studies on Ephemeroptera III. Coloburiscus humeralis (Walker): morphology and anatomy of the winged stages. *New Zealand Journal of Science*, *8*, 398-415.

Wood, V., & Pawson, E. (2008). The Banks Peninsula forests and Akaroa cocksfoot: explaining a New Zealand forest transition. *Environment and History*, *14*(4), 449-468.

Wright, S. (1943). Isolation by distance. Genetics. 31. 39-59.

Wright, S. (1965). The interpretation of population structure by F-statistics with special regard to systems of mating. Evolution. 19. 395-420.

Young, B. A., Schmidt, D. J., & Sheldon, F. (2012). Small-scale patterns of genetic variation in a headwater specialist mayfly: No influence of selective forest harvesting on diversity. *Austral Ecology*.

Zera, A. J., & Denno, R. F. (1997). Physiology and ecology of dispersal polymorphism in insects. *Annual review of entomology*, *42*(1), 207-230.